DOCUMENT RESUME

ED 131 043 95 SP 010 603

AUTHOR Levitov, Betty, Ed.
TITLE Licensing and Accreditation in Education: The Law and the State Interest.
INSTITUTION Study Commission on Undergraduate Education and the Education of Teachers, Lincoln, Nebr.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Teacher Corps.
PUB DATE 76
NOTE 173p.
AVAILABLE FROM Nebraska Curriculum Development Center, Andrews Hall, University of Nebraska, Lincoln, Nebraska 68588 (No price quoted)

EDRS PRICE MF-$0.83 HC-$8.69 Plus Postage.
DESCRIPTORS Academic Standards; *Accreditation (Institutions); *Educational Legislation; Employment Qualifications; Equal Opportunities (Jobs); Federal Legislation; *Performance Based Teacher Education; State Legislation; *State Licensing Boards; Summative Evaluation; *Teacher Certification; *Teacher Education; Teacher Employment; Teacher Qualifications

ABSTRACT

The legal implications of competency-based teacher education were the subject of a conference sponsored by the Multi-State Consortium on Performance Based Teacher Education. Representatives from state departments of education and their legal advisors met with a panel of consultants to discuss licensing and accrediting of teachers. The issues raised were: (1) What does the litigation say about new directions in licensing teachers and accrediting teacher education? (2) What does research say about what can be done to provide competent practitioners for classrooms and school leadership roles? The eight essays in this book evolved from the groundwork laid at this conference. The following subjects are examined: "The Law, the Courts, and Teacher Credentialling Reform," by Michael A. Rebell, a lawyer specializing in education law; "Techniques and Criteria for Designing and Selecting Instruments for Assessing Teachers," by Paul S. Pottinger, a clinical and social psychologist, Senior Associate and Director of Assessment Systems at McBer and Company in Boston, Mass.; "State Interest and Teacher Education Program Approval," "General Description of the Activities Associated with Reviews of Institutions and Programs," "Character and Use of Evidence in Program Approval," and "Due Process for Institutions and Students: Implications for Periodic Reviews," by Dr. Lawrence D. Freeman, Director of Teacher Education Program Approval, Illinois Office of Education; "Research Lags and Institutional Barriers Affecting Competency-Based Teacher Education Implementation" and a summation and report on the work of the conference by Sheila Huff, Research Fellow at the Educational Policy Research Center, Syracuse Research Corporation. (JD)

Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. Nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). EDRS is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from the original.

ED131043

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

LICENSING AND ACCREDITATION IN EDUCATION:

The Law and the State Interest

Edited by BETTY LEVITOV

LICENSING AND ACCREDITATION IN EDUCATION:

The Law and the State Interest

Edited By
Betty Levitov

With Essays By
Larry Freeman
Sheila Huff
Paul Pottinger
Michael Rebell

Study Commission on Undergraduate Education
and the Education of Teachers
Paul A. Olson, Director

1976

3

The publication of this document at the University of Nebraska Printing and Duplicating Service was made possible by a grant from the U.S. Office of Education, Department of Health, Education and Welfare–Teacher Corps. The content does not necessarily reflect the position or policy of that agency, and no official endorsement of these materials should be inferred. Requests for this book should be addressed to the Nebraska Curriculum Development Center, Andrews Hall, University of Nebraska, Lincoln, Nebraska 68588.

CONTENTS

Page

PREFACE v
William Smith, Director, Teacher Corps

INTRODUCTION ix
Betty Levitov

The Law, the Courts, and Teacher Credentialling Reform 1
Michael A. Rebell

Techniques and Criteria for Designing and Selecting Instruments for Assessing Teachers 27
Paul S. Pottinger

State Interest and Teacher Education Program Approval 87
Lawrence D. Freeman

General Description of the Activities Associated with Reviews of Institutions and Programs 98
Lawrence D. Freeman

Character and Use of Evidence in Program Approval 104
Lawrence D. Freeman

Due Process for Institutions and Students: Implications for Periodic Reviews 109
Lawrence D. Freeman

Research Lags and Institutional Barriers Affecting Competency-Based Teacher Education Implementation 113
Sheila Huff

Report of the Special Legal Issues Conference of the Multi-State Consortium on Performance-Based Teacher Education 123
Sheila Huff

OTHER BOOKS 157

ACKNOWLEDGMENT

The editor wishes to thank the executive staff and members of the Multi-State Consortium on Performance-Based Teacher Education; the Consortium funded the Denver Conference on the Legal Implications of Competency Based Teacher Education and invited the involvement of the Study Commission on Undergraduate Education and the Education of Teachers in that meeting. Bill Smith and Jim Steffensen of Teacher Corps sought out funds to print the materials included in this book as a follow-up to the Denver meeting. The editor wishes to thank Bambi Olmsted of Teacher Corps for her assistance. Special appreciation is due to Paul A. Olson, Director of SCUEET, who has provided the editor with ideas, criticism and the type of inspiration that only those who know him can fully understand. Invaluable editorial assistance was offered by Jeanne Bishop, with help from Karen Hanshew, Rich Rolfes and Suzy Voigt.

PREFACE

The areas of licensing and accrediting have been difficult areas since the two concepts were invented in the late 19th century. On the one hand, the state, under the concept of "licensing" or "credentialling," claims to protect the public safety by providing the public with competent practitioners; on the other hand, spokesmen for public interest groups have often claimed that licensing, like all forms of government regulation, serves the interest of whatever establishment is licensed more than the public interest. The area of licensing is further complicated by the fact that, in many fields, licensing is attached to accreditation—that is, a degree in a field from an accredited college often means that the degree recipient automatically receives a license to practice in the field upon application to the state.

The field of accreditation has been complicated by recent litigation claiming that accreditation agencies are in violation of the antitrust acts and recent quasi-legal action of the Federal Commission on Accreditation and Eligibility calling into question the carefulness and adequacy of the accreditation actions of several of the agencies. If a license is issued on the basis of an accreditation and the accreditation unit is no good, how can the license be good? Even if the accreditation visit is careful, does that mean that the accredited institution can promise that all of those who march from its doors can do a good job?

The determining of "who can do a good job" has also become central in civil rights litigation since the Griggs case. Now the more common industrial and business jobs require job descriptions, and employers have to be able to show that the employees whom they hire actually have the skills required by the job description. They, in some cases, need to show that they have ranked candidates for employment according to objective evidence showing the degree to which each candidate is suited for the job. That, at least, became the responsibility of the employer if challenged by a minority employee who claimed prejudice in, say, a company's personnel department.

Numerous cases have been litigated successfully, in the commoner trades and occupations, under the Griggs decision and EEOC applications of it. Now Griggs is beginning to be applied to jobs traditionally regarded as 'more sophisticated' such as the school principal's job or the teacher's job.

All of this comes to a head in the conduct of education professions. Here accreditation is attached to licensing, both licensing and accreditation are under attack, various schemes for licensing which purport to tell us which people can "do which jobs" (CBTE, PBTE, etc.) are being developed. And it is hard to claim in some cases that the "public interest" has been protected by the licensing process, particularly in minority settings where other than the English language is spoken. The courts recently ruled that children must be taught, at least in the early grades, by teachers speaking their language. And yet English language-speaking teachers have been licensed and given jobs in schools having predominantly non-English-speaking students for decades.

Recent studies show that the quality of the school staff, its fitness to the context in which it works, does make a difference to learning, to the success of desegregation policy, to drop-outs, to the emotional well-being of children—to many matters in which the state and the public have an interest.

I believe that we must meet the criteria of the Griggs decision and the EEOC guidelines in the education professions; we must have standards of employment which meet their norms—whether minorities are job applicants or not. Hence I encouraged the reports and conference summarized here.

Originally when Donald Bigelow and I asked the Study Commission on Undergraduate Education and the Education of Teachers to come into existence, we asked that it look at accreditation and licensing in the education professions. Since then the Commission has carried the job of scrutiny ahead a good deal so that we now know that some forms of licensing of teachers indifferent to Griggs or to the child's culture, some forms of accreditation indifferent to due process, will not go down at the federal or state level. I now have asked the Study Commission to secure the best information available as to:

1. How can we find out what the job of teaching is and determine in specific contexts what people are ready to do it;

2. What constitutes education for teaching competence (or for competence of any kind), and what are its pitfalls;

3. What does the law say, and what is it likely to say, about what is

requisite to a validated credential for a teacher;

4. What constitutes due process and valid state action in the accrediting of teacher education.

The conference recorded here struggles vigorously with these four questions. I commend it to the attention of state and federal officials, of people in the education professions, and of the public at large.

William Smith
Director, Teacher Corps

INTRODUCTION

Any licensing process, if it is to serve the public interest, has to be clear, has to be just and perceived to be just, and has to be validated—that is, there has to be some evidence that people who are licensed are people who can do the job. Teacher licensing, developed in this country as one of the tools through which the state exercises its police powers in protecting the general welfare, is the focus of this volume.

> A long line of court cases has established that as a general matter an individual cannot be hired and paid as a teacher or supervisor by a local school system without prior certification or licensing. Courts consistently tend to uphold the state's prerogative to exercise its police power and limit the profession of teaching to those who meet prescribed criteria. Presumably the establishment of such policies by the states, and the courts' affirmation of such policies as state responsibility, are intended to protect the general welfare. The point of the courts' decisions . . . is that licensing is a state responsibility, undertaken to insure high teaching standards at the local level—on the assumption that state officials know what constitutes good teaching at the local level. . . .
>
> But like most other occupational licensing mechanisms, those developed for the teaching profession appear to have emerged haphazardly. . . . During the past seventy-five years a system which originally placed responsibility for licensing of teachers in some three thousand local licensing authorities has been reduced to the fifty state systems and a handful of large city school districts, such as New York City and Chicago, which retain special authority. . . .
>
> The present general or qualitative requirements for licensure as a teacher (elementary or secondary) appear to be simple enough. They are threefold: (1) typically, the states have citizenship, health, age, and moral requirements; (2) all states minimally require a bachelor's degree for certification; (3) about eight states also require specialized courses in state history, state

and federal governments, agriculture and conservation (however, several of these states allow substitution of acceptable scores on proficiency exams in lieu of courses).

The quantitative requirements for licensure also appear at first sight to be straightforward. In 1973, the most common credit hour requirements for licensure among the various states were as follows: 40 to 48 semester hours in "general education" for elementary as well as secondary teachers; 18 to 24 semester hours in professional education for elementary teachers and between 18 and 21 semester hours for secondary teachers.

The seemingly straightforward nature of these requirements may be misleading, however. Most states have adopted an "approved program" approach to the licensing procedure to replace licensure by a state department transcript analysis to see if one has taken the right courses for a license. Program approval is essentially accreditation. State accreditation is, in turn, related to national accreditation.

> Final Report of the Study Commission on Undergraduate Education and the Education of Teachers, *Teacher Education in the United States: The Responsibility Gap* (Lincoln: University of Nebraska Press, 1976), pp. 102-03.

Today's accrediting systems make no guarantees as to the quality of a teacher. A program is "approved"—that is, sanctioned by various accrediting systems (national, regional, state) and indirectly, by professional societies (whose members themselves have attained professional status within these same systems). But, without more, a vast gap exists between the paper standards established by the systems and those criteria which might effectively determine if our teachers are truly competent to teach.

Not only have the teaching certification standards ignored the question of effective performance, a number of court cases in conjunction with Title VII of the Civil Rights Act of 1964 have specially held that segments of the population have in fact been illegally locked out of the education professions by present certification procedures or by examination procedures re-

lated to certification. These cases have required licensing officials to revamp their methods and assure that all job requirements are demonstrably related to attributes actually needed to perform well on the job. It appears that state licensing requiements are susceptable to challenge as follows:

1. The procedures themselves may be attacked as racially or sexually discriminatory and hence violative of equal protection and due process of law; and

2. The very foundation for licensing, the accreditation system, may be attacked as illicit state action, that is, the imposition by a body politic of a system of classification or standards which are not reasonably related to the promotion of that as part of the public welfare for which they were established.

The recent challenges to the licensing procedures and content have generally been aimed toward protecting the individual's right to equitable treatment in hiring, the public interest in the treatment of culture groups, and the denigration of the public welfare through the mistreatment of students as consumers. The *Griggs v. Duke Power Co.* decision spoke to discriminatory hiring practices. A series of cases followed in the wake of *Griggs.* Best known perhaps is the *Chance v. Board of Examiners* case, which directed the New York City Board of Examiners to revise its principal licensing processes. Decisions have seriously questioned the National Teachers' Examination, striking it down as a non-job-related or "non-validated" selection instrument. A second series of cases, such as *Wisconsin v. Yoder* and *Lau v. Nichols,* so-called "bicultural cases," focused upon group rather than individual welfare. These decisions (a) allowed parents and certain cultures to keep their children out of the public schools if those schools were somehow destructive to the cultural integrity of that group, or (b) required the schools to reflect the language and culture of the child—at least in the early grades.

An additional type of suit centered upon consumer protection. In the Bridgeport suit, a teacher education student sued the University of Bridgeport on the grounds that she did not get fair value for her money, her courses being useless and impractical. In the *Peter Doe* case, a student who claimed to be unable to read sued the school system which he had attended on the grounds that he had been passed along without being taught to read, without his parents being warned, and was consequently rendered incapable of func-

tioning properly outside the schools.

The court challenges to the licensing and accreditation procedures have resulted in some rather important, if small, changes. The accrediting agencies, for example are increasingly describing themselves as protection agencies; thus we could expect that eventually consumer protection law will come to bear on the treatment of accrediting agencies. Furthermore, the Fund for the Improvement of Post-Secondary Education requested that the American Research Institute prepare a set of consumer protection guidelines. Students have increasingly demanded a role in the licensing agencies and have succeeded in achieving the role in Texas and California. Many states have abandoned the National Teachers' Examination as a valid test for certification.

The Carnegie Commission statistics on higher education, gathered in 1970-71, and analyzed by the Study Commission, reveal that students then at least were saying that their education for the role of teacher was not adequate. Sixty-three per cent of the seniors believed that it was essential that they get a detailed grasp of a specific field, but only twenty-seven per cent have said that they received a detailed grasp. Sixty per cent said that it was essential that they get training and skills for an occupation, but twenty per cent said they received no training or skills. Yet these twenty per cent were, in most cases, licensed to teach in the public schools. In a survey done in a South Carolina city, over one half of the new teachers rated their teacher education as poor to fair. More than two thirds of the high school teachers indicated dissatisfaction with their teacher education courses. Deficiencies in their education related to their incapacity to cope with individual and group differences, to deal with student emotions, and to organize worthwhile activities for students. If the licensing process is to work, it must provide the public with teachers who see themselves as prepared to teach where they teach and whom the public also sees as competent, fit, ready. (See Study Commission Final Report, *Teacher Education in the United States: The Responsibility Gap*, [cited above] for a detailed accounting of the Carnegie survey analysis.)

There is also evidence that the licensing process is not working very well to bring into the profession people who are representative of, and responsive to, minority cultures. For example, despite the tremendous impetus of the civil rights movement, statistics suggest that the number of black teachers now is about the same as ten years ago. The National Education Association

and the Office of Civil Rights both indicated that less than three per cent of the students who needed bilingual and bicultural teachers are getting them.

* * * * * * * * * * *

The essays in this book evolved from the groundwork laid at a conference in December, 1975, on the legal implications of competency-based teacher education sponsored by the Multi-State Consortium on Performance-Based Teacher Education. At that meeting representatives from state departments of education and their legal advisors met with a panel of consultants to discuss licensing and accrediting issues. The issues raised were:

1. What does the litigation say about new directions in licensing teachers and accrediting teacher education?

and

2. What does research say about what can be done to provide competent practitioners for classrooms and school leadership roles?

Betty Levin
Study Commission on Undergraduate Education
and the Education of Teachers

THE LAW, THE COURTS, AND TEACHER CREDENTIALLING REFORM

By Michael A. Rebell

Michael Rebell, a New York attorney and partner in Rebell, Krieger, Fischbein, Olivieri, has special interest in education law. He is Special Counsel to the New York State Assembly and Adjunct Associate Professor of Law at State University of New York at Purchase, New York. Among Rebell's clients are the New York City School Boards Association, the Public Education Association, The Organization to Assure Services for Exceptional Students and the Street Academy of New York. As a consultant to the National Study Commission on Undergraduate Education and the Education of Teachers, Rebell wrote *Teacher Credentialling Reform in New York State: A Critique and a Suggestion for New Directions*, and helped to develop the chapter on the licensing and accrediting process in the Study Commission final report, *Teacher Education in The United States: The Responsibility Gap.*

THE LAW, THE COURTS, AND TEACHER CREDENTIALLING REFORM

By Michael A. Rebell

The topic I have been asked to discuss this afternoon is somewhat ironic. Gathered together in this room are representatives of education departments of 13 states which are leaders in the Competency Based Teacher Education ("CBTE") movement, a movement which is in the forefront of efforts to reform shortcomings in traditional teacher credentialling systems. But, instead of commending you on your reform efforts, my task is to outline criticisms of teacher licensing systems, including CBTE, which plaintiffs may shortly be presenting to the courts—and which the courts may be inclined to uphold. Much has been accomplished in the area of teacher licensing reform in the past few years, but, the courts may shortly be requiring that even more be done in years to come.

My remarks will stress three major themes: first, an analytic overview of CBTE from a lawyer's perspective; second, a discussion of some of the potential legal challenges to credentialling systems and to CBTE that I see in the offing; and third, a preliminary discussion of a design for a valid licensing system. Since each of these issues could be discussed for the entire conference, obviously, I will only attempt to hit the high points. I will leave some time for questions at the conclusion of my talk, and, if sufficient interest is generated, these issues can be further explored in more depth in small group discussions at a later point in the conference.

Let us turn then to the first of my three basic points, which is an overview of the legal implications of CBTE. My basic perspective on CBTE has been described in some detail in the monograph that I wrote for the Study Commission on Undergraduate Education and the Education of Teachers two years ago.[1] The starting point of my argument there was that teacher licensing, as it emerged in the 19th century, was mainly concerned with

---

[1] *Teacher Credentialling Reform in New York State: A Critique and a Suggestion for New Directions* (SCUEET, University of Nebraska, 1974).

assuring that elementary school teachers had rudimentary ability in reading, writing and arithmetic. It was not difficult at that time for a few states or local school districts to engage in examination or other processes to determine whether a potential teacher knew how to "read or write or cipher to the rule of three," as Abraham Lincoln once put it.

A dramatic acceleration in both the complexity and rigor of teacher education and teacher licensing requirements began toward the end of the 19th century as mass compulsory education became implemented throughout the country and the states felt compelled to establish explicit statewide credentialling standards that would attest to the competence of the thousands of teachers who were now being assigned to the burgeoning new school rooms. What was looked for in these credentialling standards was not merely basic ability in reading, writing and arithmetic, but also an indication of proficiency in professional techniques and a level of competence in specific subject areas. It quickly became apparent that if credentials were going to be awarded largely on the basis of proficiency in the curriculum studied at the rapidly-expanding teacher training institutions, it would make sense to allow these institutions to, in effect, become the vehicles for implementing the state's licensing authority. Thus developed the widely prevalent "approved program credentialling" system.

I will not attempt to repeat at this time the criticisms which numerous commentators have lodged against traditional approved program credentialling systems, such as the irrelevance or inadequacy of the professional courses at many teacher training institutions. (I am not sure that the elaborate modern credentialling systems have ever really proved capable of measuring any more than the same basic knowledge of "reading, writing and ciphering" that was tested by the rudimentary 19th century systems). Suffice it to say for present purposes that the CBTE reform movement of the 1970's is the latest manifestation of reform trends which have been attempting for decades to rectify the shortcomings of traditional teacher education process. The obvious need for reform was succinctly summarized by a recent statement of the New York State Education Commissioner, who, in reference to traditional teacher education programs, stated that all we can certify concerning a graduate of these programs is that he "is not intellectually inadequate and that he has some supposed interest in teaching." By implication, the Commissioner was thus indicating that we cannot certify that a graduate of a four-year program at a teacher training institution, who has been licensed by the

state, is in fact competent to teach or is any more qualified for the profession than an individual who did not undertake four years of professional training.

Now that is quite a statement. From a lawyer's point of view, the evidentiary value of such an admission by the highest education official in New York State is dramatic; the Commissioner is, in essence, admitting that the licensing standards which his state has imposed on thousands of people over the years are not capable of ensuring the state and its school children that their teachers are qualified. One obviously begins to wonder how these licensing standards can be justified or defended.

Fortunately for New York State, and of course not quite as a matter of coincidence, at the time the Commissioner made this statement, the state was already reforming its traditional approach and was moving in the direction of competency-based teacher education. I will not attempt in this short talk to undertake a description or a critique of CBTE, since I assume that everybody in this room is basically familiar with the movement's aims and methods. In a nutshell, I think the CBTE movement can be said to be attempting to meet the challenge implied in the Commissioner's statement by seeking to create a method of training teachers which will assure that the graduates of programs in teacher training institutes have been instructed and evaluated in terms of demonstrably relevant teaching "competencies" and are therefore, in fact, fully competent to teach. The award of a teacher's certificate to a graduate of a CBTE program would thus be a valid indication of a high level of professional qualification.

Interestingly, the CBTE program in New York State is being implemented in accordance with a specific timetable promulgated by the Board of Regents. A phase-in of CBTE requirements for newly-registered programs began in 1975, but the key date that interests me in that timetable is 1990, the year that the state hopes to be in a position to empirically validate the specific teaching "competencies" which the various teacher-training institutes are being required to begin to utilize at the present time. Certainly, if in 1990 the state of the art of assessment is such that discrete competencies can be empirically validated, the CBTE movement in New York and elsewhere will have lived up to its promise and will have overcome the traditional criticisms of teacher training programs. However, I am bothered by the fact that the target date is 1990—leaving immediate questions concerning the

period from 1975 until 1990. As an attorney, I must ask whether a state can impose credentialling requirements on a person who wants to teach in the public schools when, by its own admission, these credentialling requirements are presently unvalidated. Furthermore, many writers assert that even by 1990, it will be impossible to obtain satisfactory, hard, empirical validation of the discrete competencies. If that is true, there are fundamental questions as to whether CBTE, now or in the future, will be better able to withstand legal challenge than a traditional teacher-training approach.

Since traditional teacher training programs which continue to prevail in many states, as well as many newly-implemented CBTE programs, may shortly be subject to direct legal challenge, let us now turn our attention to a brief description of current legal developments which are affecting credentialling and licensing fields. Any discussion of legal challenges in the area of licensing and credentialling must, of course, begin with a reference to Title VII of the Civil Rights Act of 1964. I will not attempt at the present time to undertake a detailed analysis of the statutory provisions or specific cases decided under those provisions, with which many of you are probably familiar. For present purposes, let me just state that Title VII outlaws employment discrimination and, among other things, the statute and the regulations promulgated thereunder by the Equal Employment Opportunity Commission ("EEOC"), specify that an employer may not use any type of entry examination, whether it is a written pen and pencil test or any other type of selection system, unless that system is shown to be "job-related."

This requirement for a clear showing of "job-relatedness" has led to a dramatic increase in litigation in a wide variety of licensing areas, including the education field. I have personally been involved with the *Chance*[2] case in New York City where the federal district court invalidated the licensing examination system for principals and other supervisors. There have also been a number of cases which have successfully challenged the use of the National Teachers Examination as a job selection device.[3] In addition, the dozens of

---

[2] *Chance v. Board of Examiners* 330 F. Supp. 203 (S.D.N.Y., 1971), aff'd 458 F. 2d 1167 (2d Cir., 1972).

[3] See, *e.g. Baker v. Columbus Municipal School District* 329 F. Supp. 706 (N.D. Miss., 1971), aff'd 462 1112 (5th Cir., 1972), *United States v. Nansemond County School Board* 492 F. 2d 919 (4th Cir., 1974), *United States v. State of North Carolina* Civil No. 4476 (E.D. N. Car., 1975).

litigations brought under Title VII against other public agencies have established precedents which are of direct relevance to any analysis of future challenges to teacher credentialling or licensing systems.

Two recent decisions of the United States Supreme Court have upheld and enhanced the impact of the EEOC regulations in this area.[4] The court in these cases (which specifically involved private industry, but which, since the 1972 amendment to Title VII, apply also to school systems and other state agencies) upheld the applicability of the specific guidelines that were issued by the EEOC. This was somewhat extraordinary, because these standards were far more detailed and exacting than the general anti-discrimination pronouncements enacted by the Congress in the Title VII Statute and because these guidelines were not full federal regulations which had been submitted for public scrutiny and comment as required by the Administrative Procedure Act. But under the Supreme Court decisions, the EEOC guidelines have virtually been given the effect of law.

It is important for these reasons for us to briefly consider the precise content of the EEOC guidelines. Basically, the guidelines state that no written test or any other employment selection device may be used unless it is "validated." Validation under the EEOC guidelines is described under two main headings: "predictive validation" and "content validation." Predictive validation is the higher standard, which is normally required of all employees. Stated simply, it means that if an employer utilizes, for example, a written test as a basis for hiring, he must demonstrate that such a test is a valid indicator of actual competence on the job. He must show that those who pass the test and those who receive higher grades on the test perform better on the job than those who fail the test or obtain lower grades. In other words, he must empirically demonstrate that this test validly predicts competence on the job. To establish such predictive validation obviously is a very complicated, expensive, and time-consuming process.

Although predictive validation is the preferred standard and is required generally in cases involving private companies, in many of the public sector cases which have arisen, the courts have invoked a clause in the EEOC guide-

---

[4] *Griggs v. Duke Power Company* 401 U.S. 424 (1971); *Albermarle Paper Co. v. Moody* 95 S. Ct. 2362 (1975).

lines which permits the use of "content validation" where predictive validation is presently "not feasible." Content validation aims at the same general result as predictive validation, i.e., assurance that a test is an accurate measure of competence on the job. But, instead of requiring empirical predictive correlations, content validation requires the employer to show that the content of his test, on its face, is rationally related to the specific descriptions of the job which the employer has promulgated. For example, if a test is given for a position such as speech teaching, wherein most of the person's performance will depend on his verbal facility, a test which has one question involving verbal facility and 99 questions asking for familiarity with famous authors of English literature would not satisfy content validation requirements because the person is being tested for something that is not related to the actual demands of the job.

As I mentioned earlier, in a number of cases courts have applied the EEOC guidelines in striking down certain licensing standards, including those previously applicable to all principals and supervisors in the City of New York. Technically, the *Chance* case and many of the other early litigations against public employers—those which are most relevant to us—did not arise under Title VII because it was not until 1972 that Title VII was extended to cover state and local governmental agencies. The law formerly applied only to private companies. Interestingly, however, in the pre-1972 cases, although the challenge would initially be brought on general equal protection grounds under the Fourteenth Amendment, the plaintiffs would then "indirectly" refer to the EEOC guidelines. They would argue that, although technically the guidelines are not binding in the particular case, the EEOC, the expert agency in the employment discrimination area, has devised methods for assessing discrimination by employers, and certainly, public employers cannot be held to a lower standard of non-discriminatory performance under the federal constitution. Thus, the EEOC guidelines came to be "indirectly applied" in public employer cases and this undoubtedly provided a strong motivation for Congress to formally amend Title VII to specifically include public employers under the Act.

It is also important to note that technically both Title VII and the EEOC guidelines specifically apply only to certain defined minority groups, such as Blacks, Hispanics and women, who have been victims of discrimination. In order for plaintiffs to prevail in a Title VII case they must first show that the examination or licensing system has a "discriminatory impact" on

the minority group which constitutes the plaintiff class. That is, they must prove that Blacks, for example, are failing a licensing examination, or are being denied credentials, in a larger proportion than are Whites who are applying for the same certification. In many of the traditional written pencil and paper examinations, such as the New York City licensing examinations, such discriminatory impact has not been difficult to establish because the relevant statistics have indicated that minority groups do, in fact, fail in disproportionate numbers. Once this showing of "discriminatory impact" has been made, the burden of proof shifts to the defendant, the public agency, which is then required to show that even though this exam has a discriminatory impact, the court should permit its continuation because the test is job related and there is no alternative examination available which would not have a discriminatory impact on a minority group. Obviously, establishing such job relatedness imposes a heavy burden on an employer, especially when the rigorous predictive and content validation standards of the EEOC are the applicable guidelines.

Althcugh plaintiffs have tended to prevail in many challenges to written licensing examinations where it could clearly be shown that minority groups were failing in disproportionate numbers, for our present purposes at this conference, we must consider whether a suit against state-wide credentialling systems which rely on the program approval or CBTE approach would also have a reasonable chance of succeeding, especially if, as in many cases, few minority candidates apply for credentialling or discriminatory impact cannot easily be shown. (In contrast with a written licensing examination where one can easily compile statistics on the number of minority applicants who sat for the test and the number who received passing grades, the typical approved program credentialling system raises substantial problems such as the definition of the appropriate pool of applicants. Does it consist of those who initially applied to the teacher-training institute as freshmen and/or those who reached a certain stage in the professional training process? Is a state responsible for admission and testing processes at colleges in its state or at colleges any where in the country whose programs are reciprocally recognized?)[5]

---

[5] Some available college entrance and college graduation statistics may support an attempt to establish a showing of discriminatory impact in credentialling. See, *e.g.*, White and Francis, "Title VII and the Masters of Reality" Geo. Wash. L. Rev. (forthcoming).

If discriminatory impact cannot be established, Title VII, which is explicitly directed toward problems of employment discrimination against minority groups, would not directly apply. However, the fact that for the past five or six years, the courts have involved themselves in detailed analysis of reasonable standards for assessing job relatedness and have become educated in psychometric approaches to assessing job competence is highly significant. Given this background, one cannot assume that the courts would refuse to consider a strong attack against an irrational licensing system merely because the plaintiff is not a member of a minority group or has not statistically established discriminatory impact. Ten or fifteen years ago, a case attacking the irrationality of licensing systems probably would have been dismissed out of hand because the accepted legal doctrines of due-process and equal protection would not permit a court to closely examine the substantive licensing standards that state officials, who are presumed to be experts in the field, had established. But the EEOC has now given explicit tools which enable them to intelligently delve into these questions and to require state officials to defend the substance of their programs.

An additional legal development of recent years which lends further credence to the possibility of a successful attack on an irrational state credentialling system is the trend toward "re-vitalization" of the classical "rational relationship" standard for invoking the equal protection clause of the Fourteenth Amendment to the federal Constitution. Traditionally, the Supreme Court, especially the Warren Court, made a distinction in its equal protection decisions between those types of claims which they would examine with "strict scrutiny" (generally cases involving certain suspect categories of victims such as racial minority groups or certain "fundamental interests" such as voting rights) and all other cases which the court would review in perfunctory fashion. In a strict scrutiny case, the state would have to show that the challenged law or practice was absolutely necessary, that there was no other way to accomplish the result. Since there almost always is such a possible alternative, almost every case which received "strict scrutiny" resulted in victory for the plaintiffs. By way of contrast, in cases not receiving "strict scrutiny," the courts would invalidate a state statute or state action only if there was no possible rational relationship shown between the statute or action and any conceivable legitimate goal or purpose. If the case was analyzed under this "rational relationship" rubric, the plaintiffs generally would lose because almost any state action will have some plausible rational relationship to the apparent purpose of the statute at issue or to whatever

purpose its defenders could conceive of at the time the case was argued. In these cases, the courts would not consider other methods of accomplishing the purpose which might be less detrimental to the interests of the plaintiffs or other affected individuals.

In recent years, this strict dichotomy between strict scrutiny and rational relationship analysis has begun to break down. In the *Chance* case, for example, as well as in the *Armstead*[6] case involving use of Graduate Record Examinations as a qualification for teacher employment, the courts refused to apply the strict scrutiny standard (probably because they were bothered by the fact that the discrimination shown was not intentional). However, in applying the alternative "rational relationship" text, these courts undertook a substantive rather than a perfunctory analysis and held that the examination requirements were not substantially related to the purpose for which they exist. In other words, these courts "re-vitalized" the traditional rational relationship approach and gave it some bite.

These two licensing decisions are not isolated phenomena. Interestingly, the Supreme Court itself seems to be moving in the direction of breaking down the rigorous dichotomy in its equal protection analysis and appears to be creating a middle ground which permits substantive analysis of situations which do not involve the classical racial discrimination or other fundamental interest categories. Without exploring the issue in any great detail, I would merely point out that the Supreme Court's major decisions in this area involve sex discrimination cases;[7] one gets the impression that the court is not eager to declare sex discrimination one of its "strict scrutiny" areas because, then as I noted above, almost any case that comes before the courts will automatically result in a victory for the plaintiffs, but, on the other hand the court realizes that the rising tide of sex discrimination problems cannot lightly be set aside. This "middle ground" of a revitalized rational relationship analysis which the Supreme Court is creating in the sex discrimination cases dovetails with other trends in the licensing cases and indicates a strong

---

[6] *Armstead v. Starkville Municipal Separate School District* 325 F. Supp. 560 (W.D. Miss., 1971), aff'd 461 F. 2d 276 (5th Cir., 1972).

[7] See, *e.g. Reed v. Reed* 404 U.S. 71 (1971), *Frontiero v. Richardson* 411 U.S. 677 (1973), *Weinberger v. Wiesenfeld* 420 U.S. 636 (1975).

possibility that a valid challenge to an irrational credentialling system will be upheld by the courts, even if no racial discrimination is shown.

A good illustration of these changing judicial trends is provided by Prof. Gerald Gunther in an article which appeared in the *Harvard Law Review* several years ago.[8] Gunther referred to the case of *Williamson v. Lee Optical Co.*[9] a licensing-type case from about twenty years ago where the Supreme Court invoked the traditional rational relationship test and refused to substantively examine the state restrictions on the opticians' profession which were at issue in that case. Mr. Gunther indicated under the emerging new equal protection directions, a statute of this type, and the legislature's stated or probable reasons for passing it, would be more closely examined by the court and it might not so easily pass muster as it did twenty years ago.

In addition to a Title VII or an equal protection suit based on the revitalized rational relationship test, a number of other legal bases for potential challenges to credentialling systems should be mentioned. The first would be a "substantive" due process claim. As the lawyers in the room are aware, substantive due process is a doctrine that had been considered dead and buried for the past few decades, since the Roosevelt Supreme Court reversed its predecessors' attempts to scuttle the New Deal by reference to the businessman's "substantive due process" right not to suffer government interference with his operations. However, a number of recent privacy cases, including *Row v. Wade*,[10] the abortion decision, have begun to resuscitate the substantive due process concept. Interestingly, these privacy decisions have cited early education cases which upheld the individual's right to pursue a teaching career without unreasonable governmental interference[11] as precedent for the continued viability of substantive due process in a non-business situation. Also relevant in this regard are the alien discrimination

---

[8] "Foreword: In Search of Evolving Doctrine on a Changing Court: A Model for a Newer Equal Protection" 86 *Harv. L. Rev.* 1 (1972).

[9] 348 U.S. 483 (1955).

[10] 410 U.S. 113 (1973).

[11] Such as *Meyer v. Nebraska* 262 U.S. 390 (1923). (Statute prohibiting teaching of modern languages invalidated.)

cases such as in re *Griffiths*,[12] a case which involved an alien lawyer who was barred from obtaining a license to practice his profession by Connecticut law. The Supreme Court held there that the state's restrictions were discriminatory and unconstitutional; in the course of its opinion the court discussed in some detail the purposes and substance of the challenged licensing restrictions and rejected the state's explanations and justifications for them. In sum, this growing trend toward a re-vitalization of the substantive due process concept may lead a court to seriously consider a future claim of a plaintiff who alleges that he has been denied his substantive right to pursue a teaching career by an irrational state credentialling system.

A further, and highly likely, basis for legal challenges to a CBTE system is the doctrine of illegal delegation of legislative authority to administrative officials. Delegation was in the background of the recent opinion of the Texas Attorney General which, as most of you are aware, has led to a profound restructuring of the CBTE program in that state.[13] Although that opinion was specifically based on an analysis of certain Texas statutes which the Attorney General ruled did not permit the State Commissioner to mandate that all teacher education programs in the state be competency-based, the framing of the issues in that situation leads to the broader question of whether a fundamental policy shift such as CBTE may be mandated by an administrative agency in the absence of detailed legislative authorization.

A fifth area of possible legal challenge might arise under the First Amendment to the federal Constitution. Conceivably, a school that, for whatever reason, does not want to accept a mandatory CBTE directive from a state education department might claim that mandatory CBTE inhibits the ability of the school and its individual faculty members to teach in the manner that they believe best suits their needs, their values and their purposes.

The final type of possible challenge which I foresee is a consumer protection suit, which might be brought by two separate categories of "consumers": college students as consumers and parents of public school children as consumers. Paul Olson has already mentioned suits which have been filed

---

[12] 413 U.S. 717 (1973).

[13] Opinion H-197, January 4, 1974.

by college students who claim they have been defrauded in enrolling for courses which do not teach what they promise to teach and which do not prepare students to be competent professionals in their field. Parents of public school children may have standing to challenge the credentialling system of the state if teachers who have been officially certified as competent to teach, in fact, prove incapable of meeting their children's education needs.

In sum, then, those administrators who are responsible for the administration of state credentialling systems should be aware of these six potential types of legal challenges and, to the extent possible, should attempt to correct present shortcomings and thereby ward off such attacks. Of course, in assessing the likelihood that suits brought under any of the above six legal theories would succeed in invalidating state credentialling systems, we should consider the fact that the CBTE movement is, in general, an attempt to improve the deficiencies of traditional credentialling systems. One would expect that courts would be favorably impressed by a state's honest attempts to undertake reform. On the other hand, as I discussed earlier, a significant point about CBTE is that in criticizing traditional teacher training programs, CBTE raises fundamental questions which may boomerang and lead courts to hold CBTE programs to the demanding standards which they themselves have now established.

Since I have undoubtedly raised troublesome apprehensions in your mind concerning possible legal challenges you may face, in all fairness, I should utilize the remaining time which is available to offer a few suggestions which might aid you in defending such lawsuits. Let us then consider briefly some of the elements which would appear to be essential to the design of a valid licensing system. I will assume that the standard with which we should be concerned in this endeavor is the content validation standard. Predictive validation, that is, empirical demonstration that credentialling or licensing requirements are directly correlated to specific on-the-job performance needs, is a standard which public employers do not yet seem capable of achieving, certainly not in the teacher education area. The courts have indicated that, for the moment, until predictive validation is "feasible," they will generally accept content validation in licensing cases. Our immediate question then is, until such time, if ever, that we obtain accurate empirical assessment devices, what interim validation methods will be acceptable to the courts?

In describing the elements of a valid licensing system, the courts in

Chance, Nansemond, and other such cases, have consistently emphasized the need for a thorough, exhaustive job description. This is really the key to acceptable content validation. If you do not know what the job you're testing for really consists of, how can you purport to construct a test, or a credentialling requirement, that will certify qualification to perform that job? The judges in the above cases had little difficulty, once the issue was clarified in this manner, in striking down licensing systems when the defendants could not show that a serious analysis of the job's duties had recently been undertaken. In the *Chance* case, for example, the Board of Examiners were unable to show any organized job analysis process, and it appeared that their job descriptions for principalship positions were subjectively established by central administrators who would then consult, in *ad hoc* fashion, a number of outside experts whose qualifications or relationship to the system were unclear and whose views in any event were not solicited in any systematic manner. Three years later, when the same New York City Board of Examiners were again put on the stand to defend an attack against their teacher licensing system,[14] they were more sophisticated. This time, they came in with a huge stack of documents purporting to reflect extensive empirical investigations of what teachers actually do. One might say that in *Rubinos*, the defendants at least met the minimal responsibility of establishing their good faith and seriousness of purpose.

But I think the courts are going to require a lot more than such minimal good faith. Based on the United States Supreme Court's opinions in the *Griggs* and *Albermarle* cases, cited above, as well as my personal experience as the attorney for the *amicus curiae* community school boards in the *Chance*[15] litigation, I would emphasize the following key components of a

---

[14] *Rubinos v. Board of Examiners*, Civ. No. 74/2240 (S.D.N.Y. filed May 24, 1974). A decision is still pending in this case.

[15] As discussed above, the *Chance* case involved a civil rights attack on the examination system used for licensing all principals and supervisors in the New York City school system. The federal district court declared the examination system unconstitutional five years ago for the reasons discussed above. We have spent the last five years in New York City attempting to draw up new standards that would be acceptable to the court. Thus, *Chance* provides an ideal case study for our purposes because to my knowledge it is the only case in the educational licensing area where a new system had to be created under direct court supervision. Some of the points to be discussed in the pages

valid job analysis. First, the specific job areas should be clearly defined, but at the same time, the specificity should not be atomistic. For example, in the *Chance* situation, the consultant's report identified approximately 19 major duty areas and well over 100 specific tasks for the job of high school principal and recommended these elements as basis for the job description. The large number of task areas in this report (like the hundreds of discrete teaching competencies which have been emphasized in several CBTE competency catalogues) clearly appear to be unmanageable, while the suggested major duty areas seemed too vaguely defined to be useful.

Secondly, a representative cross section of persons who have an interest in the selection or evaluation process must have input into the job analysis. This means, for example, that the job cannot be defined exclusively from the perspective of present job holders because, especially with current pressures for upgrading and updating job responsibilities, newly-hired personnel may be expected to perform different functions or to emphasize different approaches to the job. In other words, present incumbents may not necessarily be performing correctly, or in accordance with the perspectives that contemporary policy makers wish to stress. Thus, the job description must contain a sense of direction. It cannot be a status quo report on how the job is presently being performed. Furthermore, in order to satisfy the courts, especially in a discrimination situation, administrators must be sure to include input from minority applicants or women who have been prevented from obtaining this position in the past.

In this regard it would also seem that an essential element of the job description is the perceptions of those who have a role in policy-making and defining the relevant accountability standards. Interviewing only the traditional "employer" may not satisfy this criterion. For example, the 32 community school boards in New York City have recently been vested with employment and job description responsibilities that formerly were exclusively lodged with the Board of Examiners and the central authorities, and their views therefore are important. CBTE advocates have recognized the

---

which follow emerge from criticisms which the plaintiffs in *Chance* have lodged against a job analysis prepared for the Board of Education by a professional consulting firm. (The board itself apparently has concurred in some of these views, since it recently indicated to the court that it may not adopt the evaluators' recommendations.)

need for increased participation of community representatives in the consortia that define competencies, and I think such a reflection of local needs is essential in any job description that purports to relate to actual local job requirements in a large city or a state. Otherwise significant urban-rural or cultural differences will not be reflected in the job description.

A third factor, discussed by the Supreme Court in its *Albermarle* decision, is the importance, in many cases, of having independent outside consultants undertake the job analysis. If in-house administrators who, so to speak, have an investment in the outcome monopolize the process, the courts may be suspicious.

Preparation of an adequate job description is the initial stage of the licensing process. The next step is the construction of the test or statement of standards, based on the specific job description information. The major concern at this stage is to be sure that there is a significant correlation between the test or the licensing standard and the detailed job description—that is what content validation is all about. This suggestion, of course, seems rather elementary, but in the *Albermarle* case the Supreme Court noted that although acceptable evidence was put forward validating three "job lines" at issue in the case, the defendants sought to utilize that limited information as a justification for implementing tests for eight job lines. The defendant company claimed that these job lines were very similar and overlapped, but the court held that because there was no acceptable analysis of the job skills in all eight lines, there was no basis for concluding that they were similar. On the basis of the decision in *Albermarle*, one might safely conclude that a 37½ per cent validation correlation is unacceptable. Whether the degree of correlation must therefore be 50 per cent, 75 per cent, or 99 per cent has not been made clear, but those who construct exams and standards have been put on notice by the court of the importance of this issue.

The next consideration in constructing a valid licensing system is the question of proportions. A job description may identify 20 basic job elements, but that does not mean that each element is necessarily as important as every other and that each should therefore receive an equal weighting in the overall requirements. Obviously, some aspects of the job are more important than others, and these proportions and priorities must be reflected in the examination or selection standards which are developed. In this connection, I would point out that it will be extremely difficult to validate absolute cut-

off requirements which preclude those who do poorly in one part of an exam from eligibility to undertake the rest of the test. Clearly, the concept of proportionality imposes heavy burdens on those who would maintain that certain skills, measured in specific quantifiable degrees, are absolutely essential to successful job performance.

Finally, once specific selection criteria are developed, the test or credentialling standards must be reasonably and consistently applied. In the *Albemarle* case, the company's employment selection system was invalidated by the U.S. Supreme Court largely because of a lack of objective application of the relevant job standards. The supervisory rating standards were applied inconsistently, with the same standards not being utilized in the same proportions by all supervisors across the board.

I appreciate the attention you have given to my remarks this afternoon, especially since I have gone well past my allotted time. I suspect that the impression I have left with you is that the courts are imposing extremely demanding validation standards. Leaving aside the technical jargon, however, I think it is fair to say that all that the courts really are asking for is reasonable adherence to basic standards of fairness and rationality; they are really only following the criteria that practitioners of CBTE have imposed on themselves. In this regard, one final point I would stress is the importance of providing flexible options, some alternative licensing mechanisms which credentialling candidates might pursue. If it appears that the licensing system is not overly rigid, that the CBTE experiment is not being imposed in a monopolistic manner and that candidates are given every reasonable chance to demonstrate their qualifications, the impression of basic fairness and rationality and the entire litigation posture will undoubtedly be enhanced.

Now, if time permits, I would be happy to answer any questions.

Question: In view of what you've said about different community input, are you suggesting that teachers be trained to teach in specific areas only—because that's a logical extension of what you're saying?

Answer: What do you mean by "specific areas?" I certainly wouldn't advocate that teachers be trained only for jobs in specific geographic areas. But if what you mean is training to specialize in certain types of problems and in dealing with certain types of people, I would probably favor such an

approach. Let me try to approach this question another way. It is my understanding that the logic of the CBTE system requires colleges to train people for specific jobs. The rationale for CBTE is its emphasis on direct correlation between skills actually needed in a classroom, and what students learn in the teacher training institutions. It seems to me that this means that if there is a substantial difference in the teaching environment in an urban ghetto or in a rural farm community, the CBTE program must account for those differences or it will not be preparing teachers who will be competent for the specific job. Have I fairly stated what CBTE claims to do?

Question: It depends on one's definition.

Question: Could you give a brief recapitulation of what the likely criteria are, or the area of the emerging middle, the "rational relationship?"

Answer: I assume you are referring to the "middle ground" between the Supreme Court's traditional dichotomy between cases involving suspect classifications or fundamental interests which receive "strict scrutiny" and other equal protection cases, which have received perfunctory analysis from the court so long as there was some "rational relationship" between the challenged practice and a legitimate state purpose.

An example of strict scrutiny analysis, which is traditionally applied to voting rights cases, is provided by the poll tax situation. Let us assume that you have a state law which requires a poll tax. There are obviously many valid, plausible justifications for poll taxes: they are a method for raising money, and for assuring that those who come out to vote have some interest in and hopefully some knowledge of the issue at stake. It is a system that has been in effect for generations, and therefore, it would seem, on its face, to be a "rational" system. But because voting rights have been defined as being an area of "fundamental interest," the Supreme Court[16] invalidated a poll tax statute which interfered with those rights even though the law might be "rationally related" to legislative state purposes such as raising taxes or motivating voter interest. Applying strict scrutiny, the poll tax was deemed not to be a necessary state action since, as the court implied, reasonable alternatives which affected the plaintiffs' interests less detrimentally, were

---

[16] See *Harper v. Virginia Board of Elections* 383 U.S. 663 (1966).

available to accomplish the state's purposes. If that case had been analyzed under the rational relationship criterion, however, the poll tax undoubtedly would have been upheld because obviously the tax did have a rational relationship to the legitimate state aims of raising taxes, motivating voters, etc.

Now under the middle ground or "re-vitalized rational relationship" approach I've been describing, the state will not necessarily have to prove that there is no alternative way of achieving a valid state purpose as in the strict scrutiny situation. Let me use the *Lee Optical* case as an example to try to bring out this distinction. As I recollect, the issue in that case was that opticians were barred from making glasses without a prescription from an optometrist or a physician. In other words, they were not permitted to examine people's eyes or even to prepare a prescription from studying an old or broken pair of glasses, unless they were licensed as optometrists or opthomologists. The Supreme Court's 1955 decision indicated that there is a "rational relationship" between the valid state aim of protecting eyes and encouraging eye examinations and the particular statute which had the effect of requiring a child who broke his glasses or even his frames to undergo an optometrist examination before the glasses or frames could be replaced. The court speculated as to a number of purposes the legislature might have had in mind in passing the statute (except, however, for the most likely real reason—the strength of the optometrist's lobby) and held that even though the laws here may on balance be "needless" or "wasteful," if the legislature could be deemed to have one or more plausible intentions, its statutory scheme would be held acceptable under the rational relationship rubric. Now, under the new middle ground standard that I have described, the court in a 1975 *Lee Optical* case would look more closely at the real purpose that the state legislature had in mind and whether the statute at issue fairly and substantially accomplished that purpose. Under such an approach, the *Lee Optical* case might well be decided the other way, if the court were to probe and find an improper state purpose (aiding the optometrists' lobby) or an irrational implementation mechanism which *heavily* burdens both the opticians and the public to accomplish a result which might easily be better accomplished by other means such as requiring eye examinations for school entry, renewal of drivers' licenses, etc.

Question: Why do the courts seem to be preoccupied with teacher licensing systems?

Answer: I hope I have not given a misimpression. Many other state licensing

systems are currently being challenged, teaching is by no means an exclusive target. Numerous cases have been brought against exams for policemen, firemen and other civil service workers, including the members of the New York City Board of Examiners themselves. For some reason, there have also been a spate of cases in various states attacking licensing requirements for undertakers and funeral directors. Teacher cases gain a certain amount of publicity because the teacher ranks are so numerous and because there is great public concern in this area. The legal standards that I have been discussing apply across the board in a great number of areas. The cases that have invalidated undertakers' examination, police exams, or whatever, provide legal precedents, pro and con, for what goes on in the teacher licensing area and vice versa.

Question: How do the Title VII requirements apply to state credentialling requirements?

Answer: The EEOC regulations specifically apply to "ability tests which are designed to measure eligibility for hire, transfer, promotion, training, referral or retention. . . ." The term "test" includes all formal, scored, quantified or standardized techniques of assessing job suitability including . . . specific qualifying or disqualifying personal history or background requirements, specific educational or work history requirements,"[17] etc. What this appears to mean is that any standard or criterion that is used to restrict employment opportunity, including a college or professional degree requirement, would be subject to the Title VII standards. Now, with a private employer, restriction on employment means declining to accept someone who applies for a specific job. One might say that credentialling is a step removed from this situation because accreditation is not synonymous with being hired for a particular job. But, I think that credentialling practices, at least if they are shown to be discriminatory, would be covered by the EEOC guidelines because they restrict entry; if you do not qualify for the license, you will be denied access to the job.

Question: Was the National Teachers Examination itself struck down in the cases you mentioned, or was it the way the test was used that was struck down?

---

[17] C.F.R. § 1607.2.

Answer: The courts invalidated the use of the NTE as an exclusive basis for rejection of applicants for teacher jobs. Specifically, the courts held that a cut-off NTE score was not a valid method for eliminating applicants since there was no showing that questions asked on the NTE had any direct relationship to actual requirements of the job. Interestingly, in many of these cases, representatives of the Educational Testing Service, which created the NTE, testified that the exam was being misused and it had never been intended to provide an exclusive basis for teacher hiring decisions.

Question: I see only three basic alternatives coming out of the *Chance* case: (1) a watered down minimal state standard for licensing to try to encompass the state as a whole; (2) the growth of a number of licensing titles that are just absolutely mind-boggling; and (3) wiping out of state licensing systems.

Answer: I think you are quite correct in your perceptions. As you know, I have been personally inclined to favor the first or the third of the possibilities you mentioned, at least on a temporary basis until validated new licensing systems are created, so I am not shocked at the implications of your statement. I am sure we differ on that. In regard to your second possibility of a proliferation of job titles, I might point out that one interesting result of the job analysis for New York City supervisors was the consultants' recommendations that many of the existing job titles be combined. However, that may be a unique situation because New York had something like a hundred different job titles for supervisors, which, of course, is absurd since, obviously, many of them must have overlapped. If your present system has two or three job titles, it is possible that a job analysis will result in breaking these down into a dozen or more.

Question: Your comment about urban-rural is a direct indication that now there is no distinction made under the present system, and if you go to a court test of the state licensing system, which I think might be interesting, that may stop some present movements that are pretty inappropriate.

Answer: Maybe. I think one other point I should mention about the *Chance* case is that the court there has already begun to move in the direction of emphasizing local orientation in licensing standards. The new system negotiated by the parties and approved by the court in that case establishes city-wide entrance examination standards which should be supplemented by the local community boards. Then, after one year of on-the-job experience, the

candidate will undergo a performance evaluation for a second stage licensure. This second stage is obviously more community-oriented.

Question: There are some people in this room that don't think directly in that line and don't accept minimal standards as a basis.

Answer: Given the present state of research and given all the inadequacies one can point to in applying the EEOC standards to most of the present licensing systems, my natural reaction is to say that until the state can come up with a solid licensing standard that meets validation criteria, it has no legal right to impose employment restrictions on people. Until you can come up with meaningful, validated requirements, you should only be allowed to apply minimal standards because a state should not attempt to declare that certain people are not qualified for the profession if it does not really know what qualification in the profession means.

To the extent that we presently lack knowledge to construct valid state-wide standards, or at least rigorous standards, I personally tend to emphasize on-the-job evaluation as being the most reasonable available assessment technique. On-the-job evaluation, of course, can only be done in a meaningful way on a local level, according to local needs and local assessments. Now I realize the calibre of evaluation undertaken by local school districts in general has been even more backwards than state-wide licensing practices. But in terms of short-run feasibility, until we reach the stage of accurate predictive validation for broader standards, it seems to me more practical to attempt to immediately buttress local on-the-job evaluation practices.

Question: Isn't on-the-job evaluation really done through a probationary period, and isn't that really an employer prerogative and not a state prerogative?

Answer: Yes, that's true, and again, my point is that until the state is in a position to apply uniform state-wide criteria that are validated in a meaningful sense, I do not see how it can interfere with a local employer. The fact that local evaluations necessarily are classroom-based and closer to actual instructional needs means that local evaluation, at least potentially, is more in tune with the aims of CBTE. I agree with you, evaluations are not currently being adequately undertaken in most communities, but my point is that immediate attention and effort should be turned toward this area, which gives

a promise of greater immediate feasibility.

Question: You talk of different standards in every district of the state. If I'm a teacher seeking a job, or an administrator, am I not being discriminated against if there are, in fact, as many different standards as there are districts or schools?

Answer: Not if the standards are clearly articulated, and are applied fairly and objectively. I know that is saying a lot. Presently (at least in theory) there is one set of standards in New York, and another set of standards in Colorado. Is that discrimination if a teacher who was trained in Colorado does not qualify under New York standards?

Question: What has been the position of the teacher unions in your licensing cases and what are the political implications of courts getting involved in these cases?

Answer: As far as the unions' position in the litigations that I've been involved with, in both the New York City supervisors' licensing case and in the teacher licensing case, the unions intervened on the side of the defendants, and fought vigorously to uphold the existing licensing system. In many ways, they fought harder than the original defendants themselves. Interestingly, in *Chance*, the case was brought against both the New York City Board of Examiners, which is a separate statutorily-established agency, and the Board of Education and its Chancellor. The Board of Education did not put up a vigorous defense and the Chancellor, in fact, refused to defend the case and even submitted an affidavit to the court saying that he personally favored the position of the plaintiffs; he implemented the licensing system because he was required to do so by state statute, but he would not personally defend the system before the court. The Board of Examiners, however, decided to defend the action to the maximum extent possible. Three or four years later, when the teacher licensing case was similarly brought against the same defendants, the Board of Education as well as the new Chancellor decided to vigorously defend the action, and that case is still pending. I don't know what was the reason for this change of position. Perhaps it was because the composition of the board and the chancellorship changed, perhaps the political climate shifted over the years, or it may be that they thought that the teacher's exam was more valid than the supervisor's exam had been.

Turning to the larger question of involvement of the courts in the

politics of education, I would first note that the court's involvement was not self-motivated. One might say that Congress invited or required them to get involved by passing Title VII and by extending it to local governmental agencies. Now inevitably, if a court rules that a licensing system is unconstitutional, its on-going supervision of the implementation of new system becomes quite extensive. The *Chance* case, as I mentioned, has been in litigation for five years, and essentially, over that period of time, the Federal District Court has become a co-administrator of the entire New York City school system. There is almost no important decision regarding supervisors that can be made in the city without the court's involvement. For example, last fall the system was faced with the prospect of massive position reductions because of budget cuts. The plaintiffs alleged to the court that the contractually-mandated seniority lay-off system was in conflict with the court's ruling in this case. They said that the court intervened, minority supervisors who presumably were prevented from being hired and from accumulating seniority by a discriminatory examination system would be the first to be laid off. So no lay-offs have been effected while the court has deliberated on this issue, and the lengthy appeals process has been completed.

On the other hand, the courts are very aware and very wary of their political role and their involvement in day-to-day administrative decisions. At every stage in the case the judges (over the past five years, there have been three of them) have constantly attempted to narrowly define the issues and to leave as much discretion as possible to the school administrators, but the nature of the situation sometimes makes this impossible. This attitude may explain why content validation has been accepted as a reasonable interim procedure by the courts. If the more demanding predictive validation standard were to be applied, the courts might feel that they would never be able to terminate their oversight role.

TECHNIQUES AND CRITERIA FOR DESIGNING AND SELECTING INSTRUMENTS FOR ASSESSING TEACHERS

By Paul S. Pottinger

Paul Pottinger is a Senior Associate and Director of Assessment Systems at McBer and Company in Boston, Massachusetts. He is an experienced clinical and social psychologist whose background includes university level teaching and significant work in the practical application of the behavioral sciences to organization, manpower, and management problems. Additionally, he has had considerable experience in the planning and management of research programs in the U.S. Department of Health, Education and Welfare.

Pottinger's special interests and expertise are in the design and implementation of assessment systems where the identification and measurement of individual competence related to work and other social roles lead to equal employment opportunities and organization effectiveness. He has successfully utilized social psychological techniques to assess individuals' educational and training needs and in individual evaluations related to selection, placement, promotion, certification and licensing.

TECHNIQUES AND CRITERIA FOR DESIGNING AND SELECTING INSTRUMENTS FOR ASSESSING TEACHERS

By Paul S. Pottinger

The Problem in Perspective

More than ever, teacher educators want to know and need to demonstrate whether or not they are accomplishing the goal of preparing people effectively for teaching. The development and use of assessment and evaluation techniques, however, have not kept pace with the need for better answers to these fundamental questions.

Three years ago, David Krathwohl (Merwin, 1973) stated:

> One can predict that performance-based teacher education (PBTE) is certain to fail to reach its ultimate objective if it continues on its present course. *This failure will be caused by the almost complete lack of attention given to the assessment of teaching competencies,* a core concept of PBTE. Only by such assessment can we achieve the goal of assuring that a teacher can indeed perform in ways that result in children learning. (Italics mine.)

Merwin went on to say in this same publication:

> Indeed, the ability of trainers to adequately define and measure competency lies at the heart of the question of whether PBTE can be successful.

Dr. Krathwohl further commented that there is a paucity of instruments available for assessing teachers. This is still true today, and it is one reason, no doubt, that I have been asked to address this group on criteria for development and selection of instruments. My comments will focus primarily on development, partially because I am aware of few relevant

instruments from which to select.

Others have written eloquently about the problems of PBTE (see especially Broudy, 1972 and Merrow, 1974). Still others have discussed issues of assessment research which provide useful criteria in the development or selection of instruments (see for example, Merwin, 1973; and Medley, Soar and Soar, 1975). I will not attempt to summarize these readings since you are no doubt familiar with them and many others already.

This presentation will:

-- briefly discuss a variety of perspectives that impinge upon assessment and performance-based teacher education.

-- suggest a conceptual model for meeting the simultaneous demands of those who have different perspectives yet are in search of a common ground for achieving practical solutions for validating teacher competencies.

-- discuss some specific assessment instruments, developed or utilized in a variety of educational and occupational settings, which serve as prototypes and models for new instrument development and selection.

Background--Social vs Scientific Impetuses for Change

The impetus to change assessment of teacher performance has come primarily from the courts, taxpayers, and educational consumers. This impetus has culminated, in New York State at least, with Commissioner Nyquist's challenge that we must assure the public that graduates of accredited teacher education institutions are more competent than non-graduates. Performance-based teacher educators are trying to instill public confidence in teachers' competence by empirically validating specific behavioral objectives. Research, theory, and development, however, have not kept pace with the social, political, and legal requirements for change that are being imposed.

While we should welcome incentives for change, we should recognize them as a mixed blessing. New York's validation goal is a frightening proposition for higher education researchers who find little support for the rigorous

research needed to meaningfully validate teacher competencies. Government and private sector funding sources as well as program developers and researchers are responding more to the social forces for change than to rigorous empirical (i.e. scientific) quests for better, more practical solutions to complex problems. PBTE efforts have reflected a reactive movement in response to simple and minimal court demands for content validity and quick answers rather than a proactive movement in response to fundamental scientific demands for better knowledge of cause and effect relationships. If we attempted to answer these more fundamental questions, the issue of predictive validity, rather than mere content validity, would receive more attention than it now does.

Thus, social pressures have led educators into attempting to clarify objectives, goals, and outcomes without first establishing a strong methodological base from which to make reasoned decisions. Such decisions are instead being made from considerations which lack the vision, concepts, and methods necessary to meet meaningful, long-range assessment goals.

Performance-based teacher education appears to be a rational approach for making program and individual certification decisions. Its efforts to date, however, have fallen quite short of meeting the needs for a system that is not only defensible in the courts but, perhaps more importantly, defensible to students, teachers, taxpayers, program developers, and others who want to know what constitutes effective teaching.

The courts are not requesting the quality research needed to answer fundamental questions of cause and effect relationships because their analyses, albeit systematic and rigorous, are aimed instead at assuring due process. That is, the court's rightful role is to demand fair processes of decision-making within a framework of what is known. The researcher's role, in contrast, is to improve upon this knowledge base rather than to assure fair and equitable use of it. As long as educators respond to the rigor and logic of due process rather than to the rigor and logic of science, PBTE will not withstand the challenge of the courts any better than standard or traditional educational processes.

One cannot ignore the social pressures for change which arise from legal, economic, political, and other social conditions of the day. But we must be clear about the differences in goals and strategies for change which these

forces create. The education researcher may be interested in validity, theory, methodology and other scientific concerns which are of little interest *per se* to taxpayers or the courts. A teacher may be primarily interested in being fairly assessed for a broad range of knowledge, skills, abilities and other characteristics that constitutes his/her competence; the teacher is appropriately more interested in comprehensive and fair assessment than in scientific issues. The education administrator must be concerned about accountability demands from accrediting organizations. The State Board of Regents is concerned with accountability demands from the courts and the public in general. And the student and his parents want more for their consumer dollar than is often demonstrated by our teaching institutions. Each perspective is legitimate but no one perspective should be allowed to dominate our strategies for solving the problems which ultimately each of us shares. Adequate criteria for selection and development of assessment instruments should keep these differing and relevant perspectives in balance. To maintain this balance, we need competent integrators who can translate the critical needs of practitioners, administrators, parents, students, and the government into practical scientific approaches to problem solving.

The Researcher's Perspective

For the researcher, a more salient problem than maintaining a balanced perspective is, frankly, defining what competence is. And that is a particularly delicate issue if one considers measuring teacher competence in terms of student outcomes. Let's look briefly at some real-world barriers to answering this question satisfactorily.

The objective of schooling, for example, becomes relevant to defining competence in terms of student outcomes. Is the objective to foster the general transition of children into adults? Is it to prepare people for the world of work? Is it to function as a socio-political-economic gatekeeper with respect to who attains further educational and career opportunities? These questions represent philosophical and ideological strata which are more properly the domain of values than of science.

Assuming there is agreement about the objectives of schooling, we would then ask what student outcomes are appropriate to these objectives. Are they primarily having a grasp of the three R's? Or do they include interpersonal competencies? Are they critical thinking and problem solving

abilities? Are they moral behaviors? Do they include the constructive and effective use of power or an orientation toward achievement? If we can answer the first set of value-laden questions, we can at least begin to answer the question of appropriate outcomes with some empirical rigor. Until we have done so, we cannot very meaningfully ask what teacher behaviors and other characteristics best enable these learning/development outcomes. Yet, little progress has been made toward developing rigorous, useful and relevant operational definitions of student outcomes. No wonder there is simply no good evidence to relate teacher behaviors to student outcomes.

The reason there is so little evidence to provide us with fundamental knowledge about what works is, of course, the sheer complexity of the problem. There are too many unknown or uncontrollable variables. Students differ in their basic behavioral repertoires, socio-economic backgrounds, motivations, interests, values, and perceptions of their environments. Teachers also differ according to these and other important dimensions. Finally, the environments within which teachers and students interact differ according to numerous physical, psychological, organizational, social, economic, and political dimensions. Medley, Soar and Soar (1975) have eloquently presented a model for conceptualizing the complexities created by these variables, and they have elaborated on the implications of these obstacles for doing research. We will refer to their four-level paradigm for conceptualizing research in the next section.

One reason for elaborating on the socio-political context of competence assessment as well as on the scientific complexity of the problem is to make two points more salient. That is, (1) we cannot meaningfully discuss criteria for developing or selecting instruments for assessment without agreeing upon the context or paradigm from which we operate together. Part II of this presentation will briefly describe the paradigm from which I am operating. Suffice it to say here that selecting and developing measures of competence is to my way of thinking, putting the cart before the horse, since there is little or no scientific evidence about what constitutes competence in the first place. (2) More systematic collection and analysis of data is required of all of us if we are to make significant progress toward the goal of empirically validating teaching competencies by 1990. We need not, by the way, agree upon philosophical and ideological issues in order to do empirical research. Also, heavy emphasis by researchers on rigorous scientific analyses of competence should not denigrate practitioners. Many practitioners have

strong convictions about what constitutes quality education, but they are unable to empirically validate these convictions. We should not belittle them for assessing student competence on a subjective basis; rather, we should tap their intuitions and covert working hypotheses to uncover clues to the causes of successful teaching.

The plea for more scientific research stems from the belief that such research is critical to the development of quality PBTE programs that are attempting large-scale change in the way they teach, assess and credential students. Moreover, the outcomes of quality assessment research could replace legal and political pressures as the "prime mover" in accomplishing the changes.

Performance-based teacher education will not progress very far unless it is able to move beyond the subjective assessment of the "laundry lists" of behaviors which proliferate the field. No matter how strongly current curriculum and assessment innovations are supported by the courts and others who demand equity and accountability, PBTE must provide scientific evidence that it works better than the *status quo*. The barriers to defining and measuring competence will not be eliminated by arguments for change and innovation no matter how strongly they side with equity, accountability or other relevant social goals. The outcomes of PBTE must be developed and evaluated systematically, rigorously and scientifically.

Most new attempts to define and measure learning outcomes according to what people can do are restricted in scope, lack rigor or poorly correlate with job and life requirements. The state of the art of assessment is in need of more conceptual rigor, systematic and comprehensive strategies for developing measures, and empirical verification of the utility of these measures.

Until we have a more comprehensive base, educators will continue to use an existing array of questionable measures based on narrow cognitive outcomes, superficial behavioral observations, and *a priori* value-laden judgments. What is required is a sophisticated technology capable of uncovering critical factors which are necessary and sufficient for competent performance.

Criteria for Assessment Instrument Development: Locating a "Rational Center"

Moving from Behaviors to their Causes

The one criterion I wish to emphasize is that teacher competency definitions and assessment instruments should comprise causally related, criterion referenced and generalizable variables which take into account complex observable and unobservable behavioral repertoires required for successful performances. Underlying this statement is intuitive, theoretical, and empirical evidence supporting the notion that competencies cannot be meaningfully defined by seemingly endless lists of specific skills, tasks and actions which ultimately fall short of real-world requirements for effective performance. In fact, the more essential characteristics for success often turn out to be broad or generalized abilities and characteristics; and they are sometimes more easily operationally defined and measured than an array of specific "subskills" which do not add up to general competence.

Most PBTE assessment techniques are based on the observation of external behaviors which are the building blocks of successful performances. But these assessment techniques tend to be reductionistic and lacking in meaning because *they fail to assess the underlying causes of these behaviors.* Usually the result is the assessment of behaviors which may have little generalizability or transferability to a variety of teaching requirements. This problem has important implications for curriculum and instruction, because observable but superficial behaviors rather than their causal underlying factors often are taught. Thus, what is actually learned, as well as what is assessed, may have little general significance in post-academic life.

This bias will arouse strong feelings in many supporters of the PBTE behavioral objectives approach to competency definition and measurement, so I will try to put this criterion into an historical as well as methodological perspective.

Again, differing perspectives of program developers must be taken into account. I hope to find some practical middle ground, not because it is safe, but because I think that recognizing the major thesis, the need to measure causal viables, can potentially lead to pragmatic solutions for finding what Merrow (1974) calls a "rational center" in PBTE.

Further Background:
Reliability and Specificity vs. Wholism and Relevance

A basic tenet of PBTE, as I understand it, comes from the Committee on PBTE in its sixteenth publication entitled, "Achieving the Potential of Performance-Based Teacher Education: Recommendations," (1974). This document states,

> The unique strength of PBTE is that it challenges all who touch it to be open about their intentions and explicit about how they will decide if their hopes are fulfilled. Ends must be made explicit; means must stand the test of relevance. The logic of the performance-based approach places a healthy stress on the use of evidence to test one's ideas and assumptions. In these considerations lie what the Committee believes to be its great potentialities (p. 29).

The interpretation and implementation of the spirit of this statement has often come from an accountability (social-economic-political) or due process (legalistic) frame of reference. These interpretations have been partially responsible for the formulation of competence definitions in terms of a plethora of professional judgments, documented by a morass of high specific but unvalidated competency statements. The lack of rigorously demonstrated relationships between competency statements (or measures) and educational outcomes has not been a serious problem for those who view PBTE as a short term pragmatic remedy to problems of accountability. Nor has this troubled those who view PBTE as a remedy to the courts' request for content valid assessment devices for licensure.

Some program developers have been led away from accounting for the richness and complexity of human behavior in their curriculum and assessment procedures by behaviorism or the "tyranny of reliability." They have fallen prey to the specific behavioral objectives approach because they believe that molecularizing global behaviors into submits is theoretically sound. Or they believe that this reductionism makes behavior more amenable to objective observation, which, *ipso facto*, solves the problem of reliability. Merwin (1973), however, pointed out three years ago that one need look only briefly at the most comprehensive collection of statements of teacher competencies available at that time, to ascertain their insufficient implica-

tions or directions for objective observation. He concluded that,

> Acting only on these competency statements, it is highly unlikely that two trainers would independently structure the same assessment procedures or that two observers would attend to the same aspects of performance (p. 12).

These comments appear to be as applicable to current competency statements as they were three years ago. That is, competency statements are not providing an adequate base for designing behaviorally specific assessment instruments.

Perhaps this failure is a blessing in disguise, for it provides us with incentives to take stock of this atomistic/reductionistic approach. Even if reliability can be increased, which some research studies have found to be the case, the *relevance* of rated behaviors seems dubious.

Implicit (and sometimes explicit) in this approach is the assumption that the acquisition of microskills and abilities adds up in some linear and simplistic fashion to overall competence. That is, highly valued generalizable skills are assumed to be mere summations of the ability to perform objectively measureable and highly specific sub-tasks.

Broudy's (1972) criticism of this assumption is useful here. The belief that the whole is merely the sum of the parts

> is a notoriously inadequate description of any human action, let alone one so complex as teaching. Teaching can, of course, be thought of as broken down into parts, but as a concrete action it is guided at every moment by a sense of its total pattern. This pattern—in swimming, reading, classifying, judging—integrates the analyzed constituents into a meaningful functional sequence, not merely a mechanically additive one. We are told, at least by some psychologists, that *after* the pattern has been sensed or felt or understood, the details can be perfected separately, but until the pattern has been discerned, drilling on the separate parts yields disappointing results (p. 3).

I think Broudy has written eloquently on this concept of the whole

vs. the parts. I shall take the liberty to quote him further. He discusses the limitations of behaviorally specific proscriptions for teaching performance in the context of three teaching styles: the didactic, heuristic, and philetic.

> Didactics refers to the impartation of knowledge by the teacher to the pupil; heuristics refers to the effort to help the pupil discover for himself either the contents of a body of knowledge or the methods of arriving at such knowledge and assessing it; philetics is merely a Greek name for love or securing rapport with pupils or, as the current jargon has it, "relating to pupils."

> Performance-based programs can accommodate didactics, which aims at more or less rote mastery of a repertoire of explicitly formulated knowledge and skill. Heuristic and philetic teaching do not lend themselves to the precise analysis, specification, and evaluation which is the presumed glory of the PBTE. Apropos of which, one might remind the namers of teaching machines that Plato and Socrates were exemplars of heuristics, not didactics.

> When a fairly reliable measure of learning is available—as it is in didactics—we can take a Skinnerian position and say, "Given teacher performance P, there will ensue pupil performance S," and we can *perhaps* ignore (for heaven alone knows what concomitant learnings take place) whatever intervenes between P and S in the minds and hearts of pupils, teachers, parents, and school boards. This is the tough line adopted by the proponents of behavioral objectives, education contractors and contractees, and the directors of the budget local, state, and national. Such toughness makes no sense in heuristic and philetic teaching, where learnings are insights and transformations of attitude for which unambiguous behavioral indices are hard to find, inasmuch as tolerance of ambiguity and lack of structure is an avowed outcome of philetics. What behavior, for example, shall we regard as criterial for a pupil's insight into his hostility to the teacher?

> Success is heuristic and philetic teaching cannot be judged by prespecified appropriate pupil behavior because such behavior

—even when we can identify it—is not manifested on demand or at a specific time. Critical thinking, the use of the imagination, warm feeling toward peers, achievement of identity cannot be inferred from one segment of behavior used as a test pattern. And what pattern shall we use as a test? Indeed, the vulnerability of general education to attack lies in the very fact that many of its benefits do not appear until fairly late in life. Our speech and reading habits, a thousand attitudes, our interests often represent the *tacit* functioning of explicit learning inputs made during school and college, but which we can no longer recall. This may help to explain why correlations between academic achievement and success in life are so low. The academic grades measured learning of items that have since been largely forgotten; functioning *now* are the residual conceptual and affective schemata, which were never tested on examinations. Nor need it be added that the life outcomes we claim for heuristic and philetic teaching are from the first contaminated by noninstructional variables, which we are never able to control adequately in our research or schooling (pp. 5-6).

A conclusion reached by Broudy in this same document is that how often one has performed a specified task or over what range may not be so important as the ability to perform a variation of the task not previously practiced. Broudy discusses this in the context of teacher training, but its implications for assessment strategies are congruent with my central thesis. Unobservable but measurable *generic and causal linking variables of* behavior are equally if not more critical than observable behaviors for defining and measuring competence. Broudy's example of a teacher whose task is to explain Boyle's law further captures the essence of this thesis:

> Suppose the prospective teacher recited the explanation of Boyle's law verbatim as it was put down in his textbook or the teacher's manual. Suppose he got all his pupils to do likewise. Would not this be proof of performance competence? Suppose, in addition, he could do all the exercises dealing with Boyle's law at the end of the chapter, and suppose most or all of his pupils could do likewise. What more definite and objective evidence of competence could one want—if that is the competence one wants? Yet it is clear that such a performance could be

brought off without either the teacher or the pupils "understanding" Boyle's law. (Indeed, many generations learned geometry in precisely this way.) As a matter of fact, a demonstration that would really satisfy us that "explaining" Boyle's law had been performed adequately would not be any specific prescheduled behavior. On the contrary, some sort of dialogue with pupils that allowed us to infer—not observe—that the basic net of concepts we call chemistry is understood by both teacher and pupils is needed. The kinds of examples and counter examples; the way pupil questions are interpreted; the cues used to set the pupil on a more profitable course; not the performance but the *state of mind* we call understanding is the crucial "product" here. No single observable behavior is likely to be sufficient proof of such adequacy, for a state of mind is not observable behavior. Skinner quite rightly doesn't worry about whether his pigeons understand what they are doing so long as they do it. If, however, the way a situation is perceived or interpreted is in any way an *important* ingredient of teaching or learning, then verbal behavior, or any other covert behavior, may not be sufficient indicators of either successful teaching or learning. In other words, performance-based teaching is in danger of capturing everything except what is most significant in many kinds of learning, viz., significance (pp. 10-11).

The implications of Broudy's comments for developing assessment instruments are many. The basic conclusions might be summarized as follows: The preoccupation with specificity, clarity and precision of behavioral objectives has left us with an operational paradigm for defining and measuring teaching competencies which is intuitively and theoretically oversimplified and invalid. Lengthy checklists of highly specific "low-inference" behavioral outcomes—even when properly sampled, factor analyzed, and interpreted—have been found empirically (though not yet legally) insufficient for validating competence. Rigorous behavioral observations fail to adequately assess either complex interactions among variables or underlying causes of competent teaching performances. Too often qualities of the mind and character, unamenable to direct observational techniques, account for the presence or absence of competent performance.

The alternative to the behavioral objectives paradigm appears to be an

operational system that leads to vague, abstract, and unmeasurable outcomes. Yet, our intuitions as well as empirical evidence tells us that competence is comprised of non-additive, non-linear, complexly interacting variables–some of which are not directly observable (see Medley, Soar and Soar, 1975 and Gage, 1976).

Causally-Related Criterion Referenced Measures: The Rational Center

What is needed, then, is a model for defining operationalizing, and measuring competence that combines the methodological rigor of the behavioral objectives approach with the richness of our intuitive, empirical, and theoretical knowledge of what constitutes competence.

One model for doing this has been developed by David McClelland, of Harvard University, and his colleagues. Its success has been demonstrated in a variety of occupational areas. As it is applied to teacher competencies, it assumes that more important than either the storage and retrieval of information or the frequency and range of tasks performed (as specified in behavioral objectives) are hosts of covert behaviors which are antecedent or causally related to successful behavior. For example: how able are people in *processing* new information for problem solving; how able are they in *integrating* this information to form new solutions; and how able are they in *implementing* these solutions.

The distinguishing characteristic of this model is that it attempts to define competence in terms of causal variables of competent behavior rather than in terms of external behaviors that are the observable manifestations or "building blocks" of competence. Such behaviors may correlate with success, but correlation or mere relatedness does not imply causation, a fact almost completely overlooked in the behavioral objectives model of PBTE. Specific behaviors may be statistically associated with teaching success, but it does not follow that these behaviors lead to success; they may have followed it. Thus, ICA's model is aimed toward defining and measuring characteristics which are not only statistically related to outcomes, but in fact, cause them.

A major assumption of this approach is that knowledge, and skills that can be defined in terms of behavioral objectives are seldom sufficient indicators of how well a person will perform as a teacher. There are many other

factors that relate to performance but are not tapped by behavioral observations. These include, for example, motivation, diagnostic listening abilities, empathy, the ability to think clearly under stress, the ability to anticipate, analyze and solve problems, the ability to be flexible in teaching style, and many others. Often these factors are intuitively obvious as critical to teaching success, but they are rarely measured effectively if at all. It is these and other variables related to complex teacher abilities that causally-related criterion measures are designed to assess.

PBTE has successfully broken away from the limited assumptions that knowledge of teaching theory alone is sufficient for developing competent teachers. But it may have swung too far to the opposite extreme of only teaching behavioral tasks. Neither the knowledge of how one should teach nor simple demonstrations of behaviors gets to the heart of competence. We must establish links between the knowledge and behaviors that constitute successful performances and the causes of these performances.

The ICA Model

What causes a teacher to utilize knowledge and behaviors effectively? ICA's operational model for answering this question would encourage one to raise new hypotheses about these causes and to validate them. For example, if what Broudy calls "insights and transformation of attitudes" or "residual and affective schemata" are important to effective utilization of behaviors, we must try to operationalize and objectively measure these causal variables. Such hypotheses might well come from observing or analyzing the heuristic and philetic aspects of teaching. We must recognize that "sensed patterns of behavior," and "perceptions and interpretations of situations," are causally related to successful performances. Educational researchers must begin to operationalize their hypotheses about these causes and see if their measures differentiate superior performers from average or poor performers. These measures should be objective, reliable, and amenable to statistical analyses, so they can be construct validated and empirically related to successful performances. (For further discussion of determining the meaning of measures through construct validations and criterion referencing, see Appendix B.)

Attention must be given measurement of what many researchers treat as "extraneous factors." Often it is these factors, considered by some to be

merely undesirable "noise" in the system, that contain rich clues to the causes of competent performances. These variables are intuitively recognized as critical aspects of success, but they have too often been treated as extraneous factors, because people don't know how to measure them. Researchers must come to grips with these critical causal components of competence. Otherwise, PBTE will fail to be meaningful to teachers and others who know that more than their repertoires of observable external behaviors are functioning in their effective teaching interventions.

If an assessment instrument validly taps a causal variable (i.e., differentiates performers on the basis of criterion measures), then the variable being measured should be taught. In the end, teaching and assessment of competence should include factors which cause effective use of both knowledge and behaviors in the classroom. For example, a teacher's ability to match teaching style with student learning style is an important teacher characteristic. It makes sense, then, to define, teach, and assess the causal variables or factors that lead teachers to make such adaptations. This approach is in contrast to merely articulating, teaching, and observing the morass of possible and necessary adaptations that will need to occur. Behavioral observations may provide evidence of whether the teacher performs specific adaptations under specific conditions, but these observed behaviors may not generalize to numerous new situations encountered in the field.

Furthermore, if the determination of teacher competence is to depend upon performances of students, we must expand the range of student learning outcomes to include cognitive, affective, and other antecedent causal variables. For example, let us assume that an agreed upon behavioral outcome for students is their ability to solve problems involving Boyle's law (borrowing from Broudy again). It may be more important to instill and assess general critical thinking and problem solving skills than merely to teach the law and demonstrate its use in limited sample problems. It is important that the student identify situations where Boyle's law is relevant as well as apply the law effectively in those situations. Moreover, the student's achievement in general problem solving or critical thinking is more important for adult life than is the knowledge and sample applications of Boyle's law out of the context of life situations calling for its use. Measures of these more generic abilities are beginning to emerge. They appear to be sensitive to important changes in students' learning; and, they measure learning outcomes which are important to fair and comprehensive determination of which teachers and

which teacher characteristics really make a difference.

In order to see the importance of these measures more clearly, I want to present the ICA model more graphically. I am taking the liberty to contrast the ICA model with the one presented by Medley, Soar and Soar (1975). This is not to denigrate their model, but to expand upon it. It is the clarity, parsimony, and conciseness with which they presented their model that allows me to do this. A brief presentation of their model and quoted explanation follows:[1]

A Simple Paradigm

It will be convenient to distinguish four different levels in the teacher's professional development at which the teacher may be assessed, as shown in Figure 1.

Level I		*Level II*		*Level III*		*Level IV*
Training Experiences	—⟩	**Teacher Performance**	—⟩	**Pupil Learning Experiences**	—⟩	**Pupil Outcomes**

Figure 1. Assessment Levels in Teacher Education

Level I refers to assessments of the *training experiences* the teacher has had: What courses has he taken? What modules has he attempted? Which ones has he mastered? Which has he bypassed on the basis of having demonstrated mastery of its objectives beforehand?

Level II refers to assessments of the *teacher's behavior* while he is attempting to fulfill the role of a teacher. What kinds of questions does he ask in interaction with pupils? How does he organize his class for instruction? How does he determine the objectives of instruction?

---

[1]This model and brief explanation is taken from Medley, Soar, and Soar, Assessment and Research in Teacher Education: Focus on PBTE, PPTE Series: No. 17, Wshington, D.C.: American Association of Colleges for Teacher Education, 1975.

Level III refers to assessment of the *behaviors of pupils* under the guidance of the teacher being assessed-assessments of the experiences they have which we all know must form the basis for any learning that takes place. What kinds of tasks do the pupils perform, in or out of class? How much time do they spend in active participation in class discussions? How often does each child receive reinforcement and for what?

Level IV refers to assessments of the *outcomes of instruction*—of those changes in behavior that it is the purpose of education to bring about. How well does the pupil read? What are his attitudes toward independent learning in adult life? What kind of a citizen does he become?

Each of these four assessment levels is represented in the diagram, joined by arrows representing lines of influence or of cause-and-effect. Thus outcomes (Level IV) are seen as influenced by, in part the result of, pupil behavior (Level III), of learning experiences the pupil has while in school; these experiences in turn are seen as at least partly determined by what his teacher does (Level II); and, finally, the way the teacher behaves—"teaches"—is affected by the experiences he has had during training (Level I). All of these things are, of course, strongly influenced by other factors not shown in the figure, e.g., community, school, pupil, and teacher characteristics.

The whole enterprise of teacher education is, of course, based on the assumption that, despite these extraneous factors, the influences that are assessed at each stage are potent enough to have an appreciable effect not only on the level immediately following but on all subsequent levels. The concept of *teacher effectiveness*, in particular, is based on the notion that pupil learning outcomes (Level IV) are affected by teacher behavior (Level II). And the justification for the very existence of teacher education is the presumption that what happens to a teacher in training (Level I) can somehow increase his effectiveness, that is, affect pupil learning outcomes (Level IV).

Implications of the Paradigm

The very existence of intermediate Levels II and III suggests that the effects of training on teacher effectiveness are subject to attenuation and are, therefore, difficult to establish. Any impact on pupil learning (Level IV) of

teacher behavior (Level II) must be achieved through pupil behaviors (Level III). Two teachers who behave identically will achieve the same outcomes only if their pupils also behave in the same way. If we attempt to relate teacher behavior to pupil learning without paying some heed to what the pupils are doing, we should no be surprised to find that the correlations tend to be low.

In the same way, the relationship between teacher education (Level I) and effective instruction (Level IV) depends on what teachers and pupils do in Levels II and III. We must train teachers to behave in such a fashion that their pupils will behave in such a fashion that the pupils will learn more!

The ICA model is presented in Figures 2, 3, and 4, where it has been juxtaposed or integrated with Medley, Soar and Soar's paradigm. The intention is to make more graphic the importance and possibilities of measuring critical causal intervening variables which are often stated as relevant to competence but rarely assessed or even operationally defined. This model reflects the notion that directly observable behaviors are insufficient indicators of successful outcomes at the pupil, teacher, and institutional levels.

Figure 2 demonstrates the notion that there are antecedent and often unobservable causal variables that affect teachers' learning experiences, teachers' performances in the field, students' learning experiences, and students' outcomes. Figure 3 demonstrates the phenomenon of intervening causal variables at the level of college-teacher interactions.[2] I have focused mainly on causal variables at the teacher and student level, but Figure 4 integrates all levels. It shows definitions, measures, and interactions of external behaviors and their antecedent and intervening causal variables. These causal variables are presented in levels A through D. Their behavioral manifestations are presented in levels I through IV. Sample measures of both types of variables are suggested at each level.

It may be the more generic causal factors of competence that should

---

[2]At the program level there are causal intervening variables that are often discussed in terms of organization and management theory, but they are rarely assessed or accounted for in curriculum and program development. Rather, the focus remains on imparting knowledge and rote behaviors without a clear understanding of how even these outcomes can be achieved most effectively.

(1) at least partially comprise competency definitions, (2) be taught in teacher education colleges, and (3) be measured in valid assessments of competence. This model would consider causal variables as legitimate learning outcomes as well as intervening variables.

Causally related criterion measures may be critical to teacher education because they provide the means for making better empirical links between teaching performances and student outcomes. This is so because these types of instruments appear to be more sensitive to short term and/or more relevant changes in student learning than standardized achievement tests.

The use of these types of measures can better clarify our understanding of what constitutes effective teaching (e.g., critical thinking, problem solving, the ability to implement effective solutions) because these measures can be made amenable to construct validation and empirical predictive validation. These validations are being demanded by researchers to understand better the

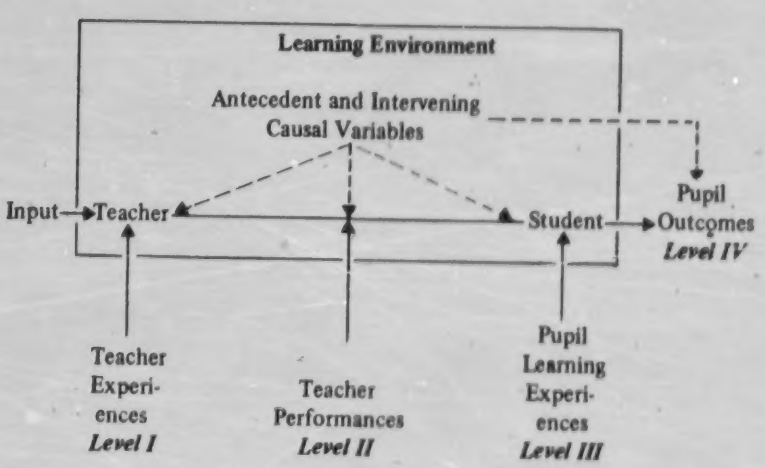

Figure 2. A Simple Model for Conceptualizing Causal Variables in the Teacher-Student Relationship

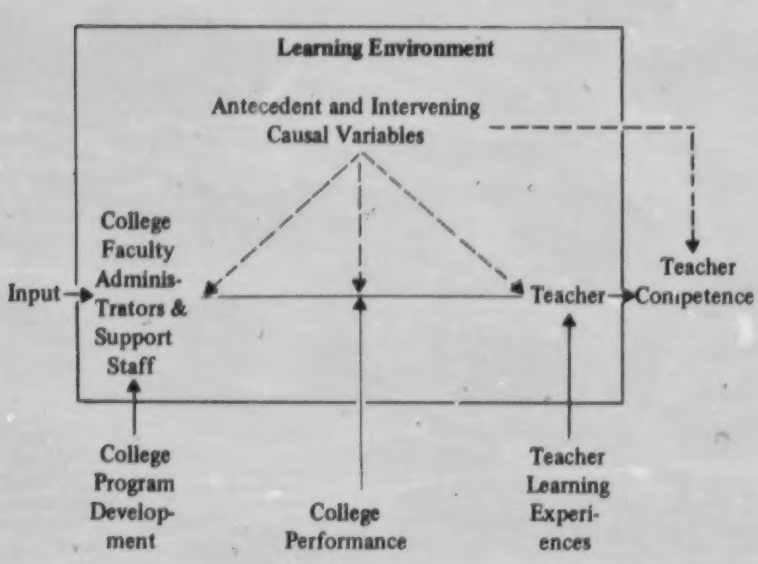

Figure 3. A Simple Model for Conceptualizing Causal Variables in the College-Teacher Relationship

meaning of measures, and they are becoming more strongly advocated by EEOC guidelines. They are also being required more frequently by our courts of law. It is because of these methodological and legal requirements for rigorous validation that behavioral checklists will ultimately have to give way to more relevant and objective measures.

Because causal related criterion measures can help clarify our understanding of what constitutes good teaching, they have important implications for curriculum, training and the modification of causal factors that underlie the behavioral performances which partially constitute effective teaching.

To assume that intervening or generic causal variables cannot be measured and are therefore unimportant is to take an unprofessional position with regard to competence definition and assessment. Researchers like McClelland, Winter, Stewart, Klemp, Kolb, Rosenthal, and others are attempting to develop valid measures of underlying causal factors.

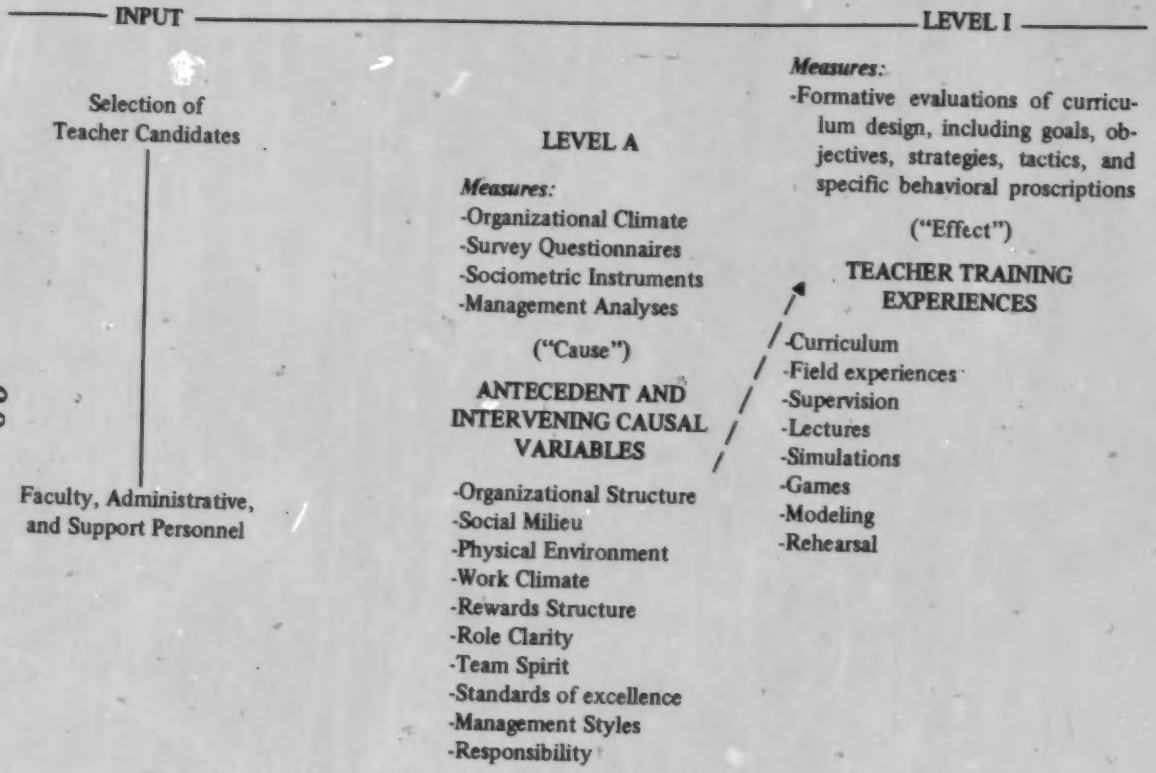

Figure 4. A Model for Conceptualizing Antecedent and Intervening Causal Variables of Institutional, Teacher, and Student Behaviors

LEVEL II

LEVEL B

Measures: (See Sec. III)
-Test of Thematic Analysis
-Cognitive Initiative
-Motivation test
-General Integrative Model
-Ego Maturity
-Behavioral Observations

("Cause")

ANTECEDENT AND INTERVENING CAUSAL VARIABLES

-Critical Thinking
-Ego Maturity
-Cognitive Integrating Skills
-Motivation
-Empathy
-Ability to Adjust Teaching Style
-Diagnostic Listening Ability

Measures:
-Check lists of Behavioral Observations
-Diaries/Journals
-Assessment Center Repertoire

("Effect")

TEACHER PERFORMANCE

-Lectures
-Counsels
-Models
-Listens
-Diagnoses
-Plans and Organizes

LEVEL C

Measures:
-Biographical Information
-Surveys
-Questionnaires
(Many of these are considered control variables to be "partialled out")

("Cause")

ANTECEDENT AND INTERVENING CAUSAL VARIABLES

-SES, Age, Sex
-Home Environment
-Time in Class (learning)
-Pupil/Teacher Ratios
-Social Milieu
-Learning Climate
-Teaching/Learning Styles
-Opportunities for Teacher Flexibility/Control

Figure 4. A Model for Conceptualizing Antecedent and Intervening Causal Variables of Institutional, Teacher, and Student Behaviors (page 2)

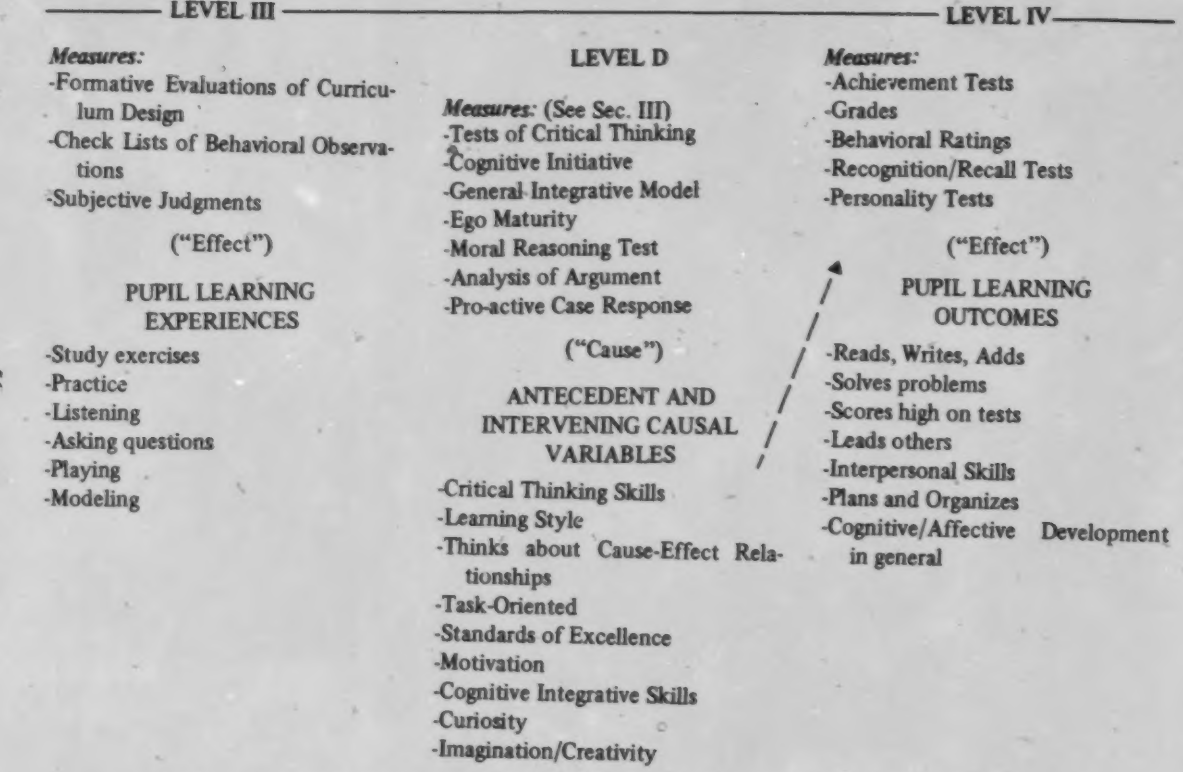

Figure 4. A Model for Conceptualizing Antecedent and Intervening Causal Variables of Institutional, Teacher, and Student Behaviors (page 3)

A sample application of the ICA model comes from research with human service workers in the State of New York. McClelland and Dailey (1974) discovered that several factors were responsible for differentiating superior human service workers from average ones. These factors included (1) faith in the ability of people to change their behavior; (2) the ability to observe and diagnose human problems; (3) the ability to arrive at realistic, non-judgmental goals with clients; (4) imagination in thinking of solutions to problems necessary to balance clients' needs with department policies; (5) persistence in pursuing solutions; (6) the ability to remain task-oriented under stress; and (7) the ability to get people to work together in a variety of *liaison* activities.[3]

In contrast, if one accepts the job description of a human service worker as a statement of behavioral objectives, the requirements for competence appear strikingly different. They describe only the behavioral manifestations of the seven factors listed above, such as "work in a counseling relationship with clients," "establish rapport," "make informed referrals when necessary," "make telephone and personal contact with agencies and services for and about clients," "have knowledge of specific community environment," "process benefits for clients within frameworks of department policies," and "make home visits." This is only a partial list. The point is there are few if any indicators of the abilities and other characteristics required to perform these tasks competently. Thus, it was more meaningful to discover and measure the generic causal variables that led to successful behaviors than it was to discover, endlessly document, and rate the specific behaviors themselves.

These findings are interesting because of the way they parallel those of education researchers. For example, Gage (1976) found that enthusiasm of teachers and teachers' acceptance of students' ideas yielded positive relationships with student achievement. Organization of teachers' presentations and teachers' use of "higher-order questions," however, were not related to student knowledge or comprehension. Yet, often, the latter two types of variables are included in competency statements. Furthermore, they are taught and assessed as if they were valid indicators of competence.

---

[3]Measures of these factors were developed and have been validated with other human service delivery personnel in the Massachusetts Civil Service Systems (McClelland and Klemp, 1975), and in Navy training centers (Klemp, personal communication).

Clearly, our intuitions about what constitutes effective teaching must be empirically verified. And we must not eliminate the verification of variables which are intuitively important but difficult to measure. Such variables may constitute the very heart of competence. Measures of these variables that are objective, construct validated and empirically linked to successful teaching might provide the means for clarifying competency definitions and the means for validating these competencies to the satisfaction of everyone concerned.

Let me summarize what is needed to establish criteria for developing and selecting relevant cost-effective and valid assessment instruments before presenting some examples of new instruments.

(1) One need is for new conceptual frameworks to guide our thinking and processes in identifying, operationalizing and measuring competence. These conceptual models must emphasize the underlying causal variables that lead to observable behaviors rather than to new ways of simply cataloging and oversimplifying the morass of complex behaviors involved in the teaching process.

(2) Better techniques are needed for developing measures which tap these broad-based, generic, or generalizable causal factors. They must emphasize the quantification of outcome criteria so that educators can rigorously and meaningfully validate these measures. They must emphasize new methods of assessing performance rather than the predominantly subjective observational checklist techniques or the passive and respondent methods (such as standard paper and pencil knowledge recall tests) now in use.

(3) Practical methods for validating new measures are necessary so that institutions of higher education understand the meaning of their assessment measures and techniques. These methods must include construct validation.

(4) New measures must be referenced to criteria which reflect requirements for teaching success if the real meaning of measures is to be relevant. These relationships must not be mere correlations between observable behaviors during assessment exercises and successful behaviors on the job, but they must reflect causal links between learning and success.

(5) Measures are needed which (a) are sensitive to short-run changes,

(b) provide useful feedback about the progress students are making toward their learning goals and (c) enable teachers and administrators to develop and evaluate better curriculum, teaching techniques, and supportive mechanisms for developing competent teachers.[4]

Prototype Measures of Liberal Arts Learning Outcomes and Other Competencies Related to Teaching

This section will describe some prototype measures which ICA has developed or utilized in liberal arts higher education programs and in other public and private sector organizations. Since they were developed or utilized in a variety of educational and occupational settings not all will have applicability to teaching competencies.

These descriptions were prepared by Dr. George Klemp, the Director of Research at the Institute for Competence Assessment as part of a proposal to develop a consortium of liberal arts competency-based colleges. The specific relevance of these measures for teacher competencies has not been articulated. Therefore, the present audience might find it meaningful to examine them in terms of understanding the types of measures ICA's empirical operational model develops or selects. Nonetheless, some of these measures are potentially quite relevant to teacher competencies, and these or similar measures should begin to be utilized in teacher education programs. ICA hopes to expand its repertoire of relevant measures for assessing competence over the next three years. This repertoire will be catalogued and disseminated. Aggregate data from their use will be analyzed and made available to the public.

By way of introduction, ICA discussed in its proposal the need for new measures which (1) are sensitive and relevant to important learning outcomes of liberal arts educators; (2) have general significance to a wide variety of

---

[4]The issue of establishing criterion levels or performance standards has not been discussed in this presentation. A brief discussion of this problem adapted from Pottinger and Klemp (1976) is presented in Appendix C. This discussion further argues against simplistic assumptions of additivity and/or linearity in assessing specific behaviors to determine overall competence.

career and life outcomes; (3) have practical utility for educators; (4) are methodologically and technically innovative, e.g., utilizing operant rather than respondent behaviors; and (5) are quantifiable and thus amenable to rigorous determination of reliability, validity, and meaning. Using these criteria ICA has developed innovative measures which attempt to answer the need for more "proactive" (operant) measurement techniques to assess the factors of process, integration and implementation.

The purpose of this section is to present information about particular instruments which have been designed to measure competency-based outcomes. A subset of these measures are discussed in depth, and data relating those measures to academic and real world outcomes are presented. For the sake of clarity, and consistent with the competency-based orientation toward outcome-relatedness, the measures described below are organized according to three outcome domains: *cognitive, effective* and *social* outcomes.

Cognitive outcomes. Measures in this domain assess characteristics purportedly measured by traditional tests of mental ability, aptitude and knowledge. The differentiating characteristic between ICA's measures and traditional tests is that ICA's measures are based on the idea that the test-taker should provide all the information necessary for adequate and appropriate response to a problem on a test, as opposed to merely selecting from a set of prepared alternative responses.

Effective outcomes. Variables measured in this domain are directly translatable to behavior patterns required beyond the world of academia. This category is derived from White's (1959) term "effectance," which means positive, goal-directed and productive interaction with and influence on the environment.

Social outcomes. These measures assess areas of interpersonal competence which often facilitate the fruition of cognitive and effective dimensions of competence in life. They take into consideration the attitudes, values and orientations toward others which moderate life goals and the means for achieving them.

Measures of Cognitive Outcomes

1. *Critical Thinking.* The ability to analyze new information and to

synthesize new concepts based on this information reflects the ability to integrate information into one's own cognitive structure. As the cognitive structure grows, so does the ability to think critically, to make a cogent argument and to reason inductively; thus, the test of Thematic Analysis is a measure of cognitive development. The test takes the form of two sets of stories which an individual is asked to compare thematically. This "thematic analysis" is scored according to twelve categories of critical thinking and a total score is derived. This scoring system is reliable, efficient and cost-effective. Each scoring category is a logical and independent dimension of critical thinking skill.

This test, developed by Winter (1973), is distinguished from other measures of critical thinking skills in that it demands the test-taker to actually produce critical arguments, rather than to simply recognize the critical elements of arguments presented to him. This instrument can be used to chart a student's progress in learning this skill. Alternative versions of the test have been developed to assess both the quality and structure of critical thinking.

Recent studies undertaken to assess the effects of the college experience upon undergraduates at Wesleyan and Harvard Universities (McClelland, 1976) show that seniors score higher than freshmen on this measure. It is important to note in this context that many so-called "cognitive" tests do not reflect the improvement in students' skill over the course of a four-year college experience. When one examines firsthand the responses to the test of Thematic Analysis, however, it is not only clear that critical thinking skills improve with college, but that the scoring system for this test is intuitively satisfying in the ground it covers.

Under an ICA contract with The Fund for the Improvement of Postsecondary Education, Alverno College began to administer the test of Thematic Analysis to incoming freshmen along with other measures, including the Watson-Glaser test of critical thinking. A chief difference between Winter's measure and the Watson-Glaser is that the latter instrument only requires students to *recognize* critical thinking (a respondent measure), while the test of Thematic Analysis requires students to demonstrate critical thinking ability (an operant measure). An analysis of the data showed that the Watson-Glaser and Winter's measure of critical thinking were somewhat correlated, but only the test of Thematic Analysis was uncorrelated with respondent measures of other unrelated abilities. Those results speak favor-

ably for Winter's measure as an uncontaminated test of critical thinking skill.

2. *Learning Styles.* A successful worker is distinguished not so much by a single set of knowledge or skills, but by the ability to adapt to and master the changing demands of one's job and career: that is, his ability to *learn.* Continuing success in a changing world requires an ability to explore new opportunities and learn from past successes and failures. Kolb's Learning Styles Inventory (1971) is a measure of individual learning styles which affect decision-making and problem-solving. The four styles, *Concrete Experiential* learning (CE), *Reflective Observation* learning (RO), *Abstract Conceptualization* learning (AC), and *Active Experimentation* learning (AE), when present in equal proportions, indicate the type of person who is able to involve himself fully, openly, and without bias in a new experience (CE), can reflect on and observe these experiences from many perspectives (RO), is able to create concepts that integrate his observations into logically sound "theories" (AC) and can use these theories to make decisions and solve problems (AE) (Kolb, 1973).

Extensive data has been collected on this measure in both college and postacademic settings (particularly the world of business). Kolb and Goldman (1973) have documented the utility of the Learning Styles Inventory for predicting major areas of undergraduate specialization and graduate school plans among M.I.T. undergraduates. The better the match between a student's learning style and the major subject area of the student's choice, the greater the tendency for students to place high importance in pursuing a career in that area, to perceive their workload as light, and to involve themselves with important peer groups, and the lesser the tendency for students to experience disaffection with their social and academic experience.

More recent work involving the analysis of administrative and technical support positions in the Division of Civil Service, Commonwealth of Massachusetts, identified "the ability to learn from experience" as a key to worker success. The Concrete Experience (CE) scale of the Learning Styles Inventory was found, in fact, to be significantly correlated with superior performance in this category of work, involving over 15 job titles (Klemp, 1976).

3. *Programmed Cases.* Based on incidents called from in-depth interviews with criterion groups, programmed cases can be developed to test

for social learning and judgment. Versions of this technique, developed for the U.S. Information Agency and the U.S. Navy, consist of a series of incidents to which several alternative responses are attached. All of the incidents pertain to a particular individual, or "case." "Distractors," or the incorrect responses, are developed with the aid of expert judges. The cases are programmed in such a way that a person with good judgment, i.e., who does not make snap, impulsive judgments, will become more accurate in his choices of the correct alternative as he proceeds through the case.

The programmed case technology has two primary uses:

- diagnostic assessment of how one uses information in making decisions about others or predicting their behaviors, and
- examination of the process by which decisions/predictions are made, including the analysis of values, biases and preconceptions that interfere with veridical impressions of others and their situations.

These programmed cases are currently being used in psychological studies at Harvard as a measure of interpersonal learning. McBer's research interest in this technology has led to applications of programmed cases in the study of prejudice.

Klemp (1975) found that people who were exposed to cases about people whose race was unlike that of the reader were less able to predict the behavior of the person in the case than readers who were exposed to same race cases. Similar studies are planned to address the prejudicial effects of socioeconomic status and sex differences on interpersonal learning.

The direct application of programmed cases, other than personnel selection, has been in assessing the skills of human relations experts in the U.S. Navy. In a pilot study (unpublished) involving selected human resource training personnel whose performance level was known, a highly significant relationship obtained between the ability to accurately predict behavior in others, as measured by the programmed cases, and performance as a trainer in human resource management.

Other measures of cognitive outcomes, in prototype form, are the following:

4. *Analysis of Argument.* A test of the ability to argue for *and* against a controversial issue, and scored for the logical presentation of argument (Stewart, 1974).

5. *Concept Formation.* A test of the ability to identify and organize similarities and differences among objects into concepts.

6. *Speed of Learning.* A test of how quickly one can learn new material selectively–that is, to remember functionally important information.

7. *Savings Score.* A test of the ability to learn new material in a particular content area–to "save" new information in an area in which the student is already well versed.

8. *Proactive Case Response.* A test of diagnosis, judgment, and problem solving that involves response to a detailed situation, or "case."

Measures of Effective Outcomes

1. *Diagnostic Listening.* The Diagnostic Listening Test consists of a taped presentation, with slides, of interviews with various individuals typical of the people one might encounter in social service work. People who take this test listen to an interview or a brief statement by a particular individual on the tape, and are then asked some questions about what has happened, what the person is really like, and what they would recommend for the person. This test requires listening, observing and judging skills which have been found necessary in human service work.

There are two subscales in this test. The Casework Subscale, consisting of 42 items, is made up of four interviews and after each of them the person taking the test is asked to answer questions and to make judgments on a multiple-choice answer sheet. The Positive Bias Subscale, consisting of 39 items, shows to test-takers three slides of clients of different sex and race with accompanying brief monologue. After each of these presentations, the test-takers are required to rate several adjectives as "does describe" or "does not describe" the client. An overall Positive Bias score is obtained by summing the number of *positive* yet *realistic* adjectives chosen. The Diagnostic Listening test measures faith in the client's ability to change, ability

to observe and diagnose human problems, ability to set realistic goals, and ability to propose imaginative solutions.

Studies of human service workers in the State of Massachusetts have verified the usefulness of the skills tapped by the Diagnostic Listening Test in identifying better workers. The format of the test instrument is similar to interview situations in which workers are involved on a day-to-day basis. Both of the two subscales correlate with effective on-the-job performance as rated by supervisory consensus (McClelland and Klemp, 1974).

Introduction to Measures 2 and 3: Much research has been accumulated by McClelland (1958, 1961), McClelland and Winter (1971), Atkinson (1958), and others that shows that thought patterns are related to important kinds of behaviors. The Exercise of Imagination is McBer's version of the Thematic Apperceptive Test (TAT) which is used to elicit thought patterns of the test-taker.

An individual taking the test is asked to write narratives to pictures. Each of these narratives addresses the following questions about the pictures: What is happening? Who are the people? What has happened in the past that has led to the situation? What is being thought? What is wanted by whom? What will happen? and What will be done? The stories are then scored, according to a *prescribed set of codes or rules*, to uncover certain patterns of thought that are expressed in the stories. These scoring codes can be applied to any written narrative which addresses the types of questions mentioned above.

The link between thoughts and behavior has been repeatedly demonstrated to be strong, as opposed to the link between attitudes and behavior. The attitude-behavior link is influenced primarily by situational factors. An attitude may represent a specific goal or objective, but such goals and objectives may change according to situational demands and constraints. However, whether a specific goal changes or not, the characteristic style with which any goal is attained is determined to a large extent by thought patterns which are relatively consistent within individuals.

2. *Achievement Motivation.* McClelland has shown in extensive research (1961) that people high in the need for achievement are practical and interested in efficiency—in short, they are good practical decion-makers.

They are independent, good at evaluating information for its practical utility, and original in the sense that they keep looking for better ways of doing things. For instance, they make good career decisions and regularly achieve greater success earlier in their careers. In a recent Harvard University longitudinal follow-up study, freshmen *n* Ach (need for achievement) scores correlated with "early success" in various fields 14 years later (McClelland, 1976).

In the world of business, studies have shown that achievement motivation is highly related to small business success, success in sales, and performance in the role of entrepreneur (McClelland and Burnham, 1976). The need for achievement, the desire to do things better than anyone else, is particularly great among scientists and others who work against a self-imposed standard of excellence. People low in achievement motivation generally do not exhibit planning or goal-setting behavior, nor do they weigh the risks they take against expected gain. The habits of behavior in such persons may not be advantageous to success in school or in many kinds of careers. But McClelland (1965) has pointed out that people can be taught to behave in ways that are reflected by the achievement motive, and so the gap between successful performance in certain academic and work settings may be effectively bridged.

3. *Self-Definition/Cognitive Initiative.* Self-definition/cognitive initiative is a general characteristic of an individual which encompasses the way one thinks about the world and himself, the way one reacts to new information, and the way one behaves. People with this competency are not only able to *think clearly*, but also to *reason* from the problem at hand to a solution, and to propose and *take effective action* on their own. Such competence is characteristic of people who think in a rational, systematic way on their own, and who can anticipate problems before they arise. In short, it might be said that people who are high in this characteristic are able on their own to see things clearly, to understand the causes of events, to reason from problem to solution, and to take effective action to solve problems. For example, the self-definition score has been quite useful in distinguishing between women who pursue careers following college and those who do not (Stewart and Winter, 1974).

A longitudinal study involving freshmen women at Alverno College begun by McBer with FIPSE funds, will track Self-Definition/Cognitive

Initiative during the four-year college experience. The preliminary data on this measure show that it is uncorrelated with other measures of college-entry knowledge, skills, and abilities. It is therefore considered to measure a unique dimension that, because of its known predictive validity regarding the success of women in careers, is a particularly important measure in a competency-based assessment system.

Other measures of effective outcomes, in prototype form, are the following:

4. *Socialized Power.* A measure of whether a person is motivated to express or increase his own power for the good of the self or for the good of others.

5. *Stage IV Power.* A recently identified measure (McClelland, 1975) of a concern for doing one's duty, that is, to be an instrument of a power which extends beyond the self.

Measures of Social Outcomes

1. *Nonverbal Sensitivity.* This test, developed by Rosenthal and his associates at Harvard University (1974), consists of 40 brief voice segments on tape, all of which have been altered to obscure the words. There are two subscales to the test: the RS Subscale, made up of voice segments that are randomly spliced and reassembled, and the CF Subscale, made up of segments which have been electronically filtered so that the words are unintelligible, but the intonation patterns remain. A sample item would consist of a speech segment followed by a question; e.g., "Does the segment represent someone helping a customer or criticizing someone else for being late?" Rosenthal has documented some promisng criterion validity for the PONS test. High scorers on this test exhibit the following characteristics:

- they represent warmer, more honest and more satisfying peer relationships;
- they have been rated by peers and/or by teachers who know them well as being generally more sensitive in interpersonal situations; and

- they were found to be functioning more effectively in the social and intellectual areas of the California Personality Inventory.

This test, which requires less than 10 minutes to administer, has been found to predict successful performance in administrative and human service jobs, which require that the worker have "empathy," or the "ability to read between the lines" in the performance of the job (Klemp, 1976). Navy personnel involved in race relations work also have been found to score higher than the general population on this test, and the personnel who are more successful on the job also score higher than their less successful counterparts.

2. *Moral Reasoning.* This test is based on the research in moral development by Lawrence Kohlberg at Harvard (1970). The test consists of a series of paragraphs which describe complex situations in which the actors are forced to choose among several moral courses of action. The task of the applicant is to write a paragraph to justify the alternative that the applicant feels is the best one on moral grounds. The essay answers are scored according to a thematic analysis developed by Kohlberg, and are interpreted according to a schema containing six levels of moral development:

Stage 1: Orientation to obedience and punishment–deference to a superior power or to trouble-avoidance.

Stage 2: Orientation to action that is satisfying to the needs of the self.

Stage 3: Orientation toward approval and to pleasing and helping others.

Stage 4: Authority and social order maintenance orientation–"doing duty" and showing respect for authority.

Stage 5: Orientation to duty defined in terms of a contract, general avoidance of violation of the rights of others, and majority will and welfare.

Stage 6: Orientation to high principle or conscience.

The conceptual categories on which the test is based have a high degree of validity as constructs.

Some recent work in the medical profession has related Kohlberg's work to the practice of physicians. High relationships exist between a physician's level of moral development and whether he will withhold or pursue treatment, the degree to which he considers the patient in the context of his family, and overall ratings of physician performance. These results show the Moral Reasoning Test to be predictive of important kinds of behavior in work which requires a good deal of value judgment. As the study of one's own values is becoming a part of what many competency-based programs wish to offer their students, Kohlberg's stage orientation to moral development is offered as an important component to this educational experience.

Other measures of social outcomes, in prototype form, are the following:

3. *Affiliation Motivation.* A measure that reflects one's concern for the compassionate quality of relationships, helpful in building good working associations with friends and colleagues.

4. *Social-Emotional Maturity.* A scoring system of written narrative built upon a stage theory of maturity (Stewart, 1973) that makes sense both theoretically and intuitively. Table 1 presents a summary of this section.

A General Integrative Model

Of the tests and measures outlined in the preceding section, none is especially useful as a diagnostic or assessment tool outside of a systematic approach to understanding the *integration* of the many skills that are required for success in life and work. The measures may be important pieces to the puzzle, but one cannot tell from pieces alone what the whole individual will look like. From the standpoint of competency-based education, it is the meaningful integration of life skills that is important as an outcome of the educational experience. The General Integrative Model is one way of expressing this value by involving several different measures in a system that can be used to assess student competence.

Table 1

Competency Based Measures and Their Developmental Status

	Prototype	Validated	Data Available in Academic Setting	Data Available in Real World Setting
Cognitive				
1. Critical Thinking		X	X	
2. Learning Styles		X	X	X
3. Programmed Cases		X		X
4. Analysis of Argument		X	X	
5. Concept Formation		X		
6. Speed of Learning	X		X	
7. Savings Score	X			
8. Proactive Case Response	X			
Effective				
1. Diagnostic Listening		X	X	X
2. Achievement Motivation		X	X	X
3. Self-Definition/Cognitive Initiative		X	X	X
4. Socialized Power		X		X
5. Stage IV Power	X			
Social				
1. Nonverbal Sensitivity		X	X	X
2. Moral Reasoning		X	X	X
3. Affiliation Motivation		X		X
4. Social-Emotional Maturity		X	X	

Such general competencies as the ability to cope with new problems, to find appropriate solutions, and to take the correct action steps can be considered in such a model. Figure 5 outlines one approximation to a systems approach that involves an integrated set of measures in a particular problem area, allows assessment at various junctures in the system for diagnostic purposes, and that also serves as a model for learning new skills through feedback in one's own performance. This particular version of the General Integrative Model requires an individual to demonstrate the following abilities:

- to observe;
- to extract relevant information;
- to analyze and integrate this information;
- to ask appropriate questions;
- to process new information in response to such questions;
- to utilize this information and one's knowledge in making sound and logical recommendations;
- to develop main and contingency plans;
- to set meaningful goals; and
- to feed back this new information into the process for better problem analysis and solutions.

This model is not a measure *per se*, but a collection of measures logically ordered, to assess problem solving skills. The progress from stage to stage in the model presents the students with subproblems to solve, e.g., what new information to seek, what conclusions to draw, and what decisions to make derived from the information gathered at a given time.

This particular model emphasizes *cognitive* skills, but other models can be developed that deal in different areas of competence. For example, the U.S. Navy, in their Human Goals Program, is striving to implement a training model that uses as input tests of achievement, affiliation, and power,

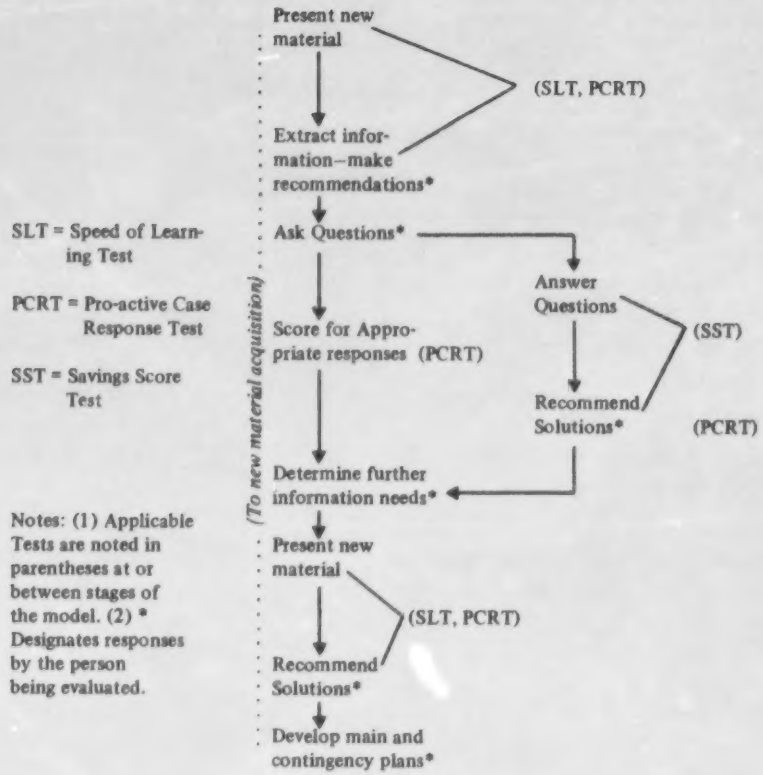
Present new material
(SLT, PCRT)
Extract information–make recommendations*
Ask Questions*
Answer Questions
(SST)
Score for Appropriate responses (PCRT)
Recommend Solutions*
(PCRT)
Determine further information needs*
Present new material
(SLT, PCRT)
Recommend Solutions*
Develop main and contingency plans*
(To new material acquisition)
SLT = Speed of Learning Test
PCRT = Pro-active Case Response Test
SST = Savings Score Test
Notes: (1) Applicable Tests are noted in parentheses at or between stages of the model. (2) * Designates responses by the person being evaluated.

Figure 5. A General Integrative Model
(One approximation)

programmed cases, learning styles, and sensitivity to nonverbal communication. By using this model, the Navy seeks not only to assess and diagnose, but to develop curriculum aimed at more effective preparation of their personnel for work.

Characteristics and Advantages of Competency-Based Measures

This section pertains particularly to the measures outlined above, but may also be considered to be the hallmark attributes of competency-based measurement in general.

1. These tests require the person being tested to be proactive, not just reactive (i.e., one has to generate responses which can be scored for their appropriateness to real life situations). Thus, the test-taker goes beyond recognizing answers out of context. In the general model, if timing of questions or recommendations is a critical aspect of problem-solving, then this time variable can be programmed into the model as well.

2. The tests are efficient since they can be given to groups as well as to individuals. Their efficiency and economy should substantially reduce the operational costs of current assessment procedures which require vast amounts of time, people and other resources.

3. These instruments foster equity in the assessment process, since they can be objectively and reliably scored according to the empirically validated coding systems. This is an important advantage since current methods of using juries, panels, or other groups to evaluate are not only inefficient and uneconomical, but are also vulnerable to all the vagaries of subjectivism.

4. The scores can be standardized with reference to criterion groups of which a student is preparing to become a part.

5. Many of these tests tap the competency of "learning how to learn" in a content area. This is one of the most important competencies people can develop because throughout their lives they will be faced with the problem of learning new things in selected areas.

6. These tests are much less threatening and anxiety-producing than traditional tests of recall or recognition, which because of their properties only contribute to the fear of failure so prominent in nontraditional students.

7. A number of variations of these tests and the General Model can be developed to add flexibility for administrators, e.g., they lend themselves to video taping, written or oral answers, individual or group testing, etc.

8. The majority of these tests have face validity. Educators and students recognize that the skills and abilities being demonstrated are applicable to general life skills.

9. Empirical and construct validation with various occupational and life skills outside of academia means that the competencies required for successful performance beyond the academic program can be established as the target of the learning process.

10. The models and tests can be validated with a variety of nonoccupation-specific populations. Some tests and models developed are noncontent -specific such that a competent person with little formal education can demonstrate competence as an analytic thinker, information processor, and a proactive initiator of appropriate solutions. The test format is easily followed and is attractive to those who are test-anxious in traditional test settings.

11. These measures can serve as pedagogical devices as well as assessment instruments, since practice in dealing with the information and component competencies necessary to solve the test problems is a direct way of learning. The instructor and student alike can easily locate and analyze weaknesses and strengths of an individual in exercising component skills. Thus, these measures can serve as diagnostic and guidance tools for supplementary curricular modules.

12. One need not take a particular course or go to a particular college in order to attain competence in the generic skills and abilities measured by these assessment tools.

APPENDIX A

BRIEF GUIDELINES FOR TEST DEVELOPMENT AND UTILIZATION

(From Pottinger and Klemp, 1976)

These guidelines are partially based upon three aspects of assessment which have been addressed elsewhere by Pottinger (1975):

- *The identification and definition of competencies* relevant to life and work outside of academia;
- *Instrumentation, techniques, and processes* of evaluation that provide reliable and valid measures of these competencies; and
- *Standardization, and/or establishment of levels* of performance necessary and sufficient for awarding credentials.

A. "New" competencies must be identified and operationally defined.

There are many outcomes of the learning experience that have greater validity than grades in school as a basis for awarding credentials. Those that have been identified have been accepted as being important and meaningful in establishing a person's competence, yet many academicians have not sought to operationalize, measure, and award credit for many of these learning outcomes. There are many other criteria than traditionally-rewarded scholastic achievement that are important as competencies in the practical work, and most of them are as yet unidentified.

B. New competencies should have general significance to a wide variety of career and life outcomes.

Competencies cannot be meaningfully defined by a seemingly endless reduction of specific skills, tasks and actions which ultimately fall short of real-world requirement for effective performance. In fact, the more essential characteristics for success often turn out to be broad or generalized abilities or characteristics which are sometimes more easily operationally defined and measured than an array of specific "subskills" which do not add up to general competence.

C. New definition of competencies and measures developed for their assessment should be easy for faculty and students to understand and use.

New competency definitions should be readily recognizable as important, and related assessment techniques and instruments should be easy for faculty and students to understand. It is necessary to guard against competency definitions and measures that are so complex, or trivial, or esoteric that students and faculty can neither understand them, nor accept them as meaningful and useful. In other words, educational goals should not be rendered unintelligible, and assessment procedures and instruments should not mystify the process of evaluating students' progress.

D. Competencies should be empirically linked to external realities.

Many educators assume that such things as the ability to master new bodies of knowledge quickly and effectively, to analyze and solve problems, to develop new skills efficiently, and to utilize knowledge are prerequisites for individuals if they are to take advantage of life's opportunities and surmount its difficulties. What is missing are the measures of these general abilities, which are related to important life outcomes. Only when we know what makes the difference between adequate and inadequate performance, based on empirical analyses of professions and other life activities, will we be able to develop or improve such measures, clarify new competencies, and establish credentials of demonstrated value.

E. The discovery of new ways of measuring abilities (competencies) is needed.

The measurement technology must be innovative and new, not just a new name for traditional procedures. Paper-and-pencil (objective) tests, due to method variance, correlate better with each other than they do with performance criteria. If postsecondary education is to break out of this closed circuit, different approaches to testing must be sought in areas such as learning, critical thinking, problem solving and other newly defined competencies. Measures of competence must require that the test taker *generate* appropriate learning outcome responses. The primary learning objective of education is not to help an individual select from among a set of predetermined alternatives. Rather, it is to enable a person to know how to reason; how to marshal evidence for or against an hypothesis; how to see similarities and differences in objects, ideas, and events; how to partial out crucial information from the trivial; and how to integrate these skills with purpose and meaning. Multiple-choice tests do not and cannot measure these abilities. And the behavioral observation/documentation approaches that are popular in experiential learning assessment do not allow these abilities to be measured

with adequate reliability or validity.

F. Standards of performance for awarding credentials should acknowledge levels of performance required for entry into roles outside of academic settings.

The establishment of criteria or standards of competence is one of the most difficult problems to be addressed. In every case, where standards of competence are determined for new or for more traditional outcomes, appropriate *levels* should be established by sufficient empirical evidence to ensure that they will not be viewed as arbitrary. Many educators are satisfied with *a priori* judgments of what skills and levels of performance are adequate. It is startling to realize how much we accept the face validity of credentials and how little we really know about (1) the correspondence between the abilities and levels of performance that these credentials represent, and (2) what is needed for adequate performance in life's tasks. We need to develop better benchmarks for evaluating the standards and the offerings of postsecondary institutions. (For further comments about the establishment of standards and levels of performance, see Appendix C.)

G. New attempts to define and assess learning outcomes should not be guided solely by attempts to make them functionally-equivalent substitutes for traditionally assessed school achievement.

Competency-based education requires a different type of evaluation from traditional programs to the extent that learning outcomes differ in significant ways. For example, learning outcomes in CBE are often defined in terms of what a person can do, not merely in terms of what one knows. Furthermore,

> . . . whereas in traditional programs evaluation is primarily linked to the credentialing process, in competency programs it is also used as a formative teaching tool. In other words, students are made aware of the criteria and standards for certification in a competency, and their progress is frequently measured so that help can be provided as necessary. Assessment that simply places students in a percentile or just discriminates between passing and failing is not adequate for competency-based programs. Formative diagnostic advice is needed—information that tells if the student is "real world" competent (Hodgkinson, 1975).

The temptation to restrict the development of new measurement instruments, techniques, and procedures, in order to achieve comparability with those that have gone before, has great political appeal for making such innovations palatable to traditionalists. However, if institutional and credential reforms are to succeed, we need to move beyond the recognized limitations of traditional assessment systems.

APPENDIX B

DETERMINING THE MEANING OF MEASURES

As background to this discussion, we have already stressed the importance of, for example, changing the focus of assessment from merely asking for recall and recognition of content to measuring how one processes and utilizes this information. If assessment techniques are to have sufficient meaning and credibility for determining if students are adequately prepared for teaching, we can no longer be satisfied with content-valid tests. Construct validation must be determined. Furthermore, PBTE proponents have stressed the importance of creating criterion-referenced measures which are predictive or reflective of real world requirements for successful teaching. The following discussion (Pottinger and Klemp, 1976) is a further elaboration on the necessity for construct validation and empirical linkage of measures to obtain maximum meaning of what is being assessed.

Messick (1975) has argued that, until measures have been construct validated, they lack the meaning essential to utilizing them as instruments of general education theory. McClelland (1973) further argues that, until construct validated measures use relevant real world events among their *criterion referents*, their value in assessing preparedness for work and life is limited. Educators have often failed to pay attention to construct validity because they "view desired behaviors as ends in themselves with little concern for the processes that produce them or for the causes of the undesired behaviors to be rectified" (Messick, p. 959). In other words, "construct validity is not usually sought for educational tests, because they are typically already considered to be valid on other grounds, namely, on the grounds of *content* validity" (Messick, p. 959).

In short, educators have traditionally been satisfied with knowing that the content of tests adequately sample a class of situations or subject matter. Messick (1975) argues that content validity does not provide an evidential basis for interpreting the meaning of test scores, and McClelland (1973) argues further that the interpreted meaning of scores that come from construct validation must be strengthened by tying these constructs directly to the world of events outside of academia.

The theoretical distinction between general education and competency-based education is that the latter requires an empirical and causal link between measurement responses and their meaning, as related to real world life outcomes. Most competency-based programs, however, merely correlate test responses with specific criterion-referenced outcomes (and many do not even do this) without discovering the underlying causes of these responses. Many educators make the mistake of thinking that if a test correlates with a behavioral criterion variable in the world of work or elsewhere outside of the academic world, one can develop competence by "teaching to the test." But this notion confuses correlation with causation, i.e., the fact that tests correlate with observable criteria may only indicate the existence of a causal intervening variable which is really responsible for behavior and which has not been measured.[5]

Clearly the mandate for competency-based postsecondary education is to identify skills and abilities that produce (cause) desired outcomes; to develop curricula aimed at the acquisition of these skills and abilities; and to design and validate measures that are sensitive to the acquisition processes and are representative of the criterion outcomes. One should not consider curriculum development apart from assessment issues and neither should be considered in the absence of identified valid performance criteria. Only when these conditions are satisfied does it make sense to "teach to the test."

The skills tapped by genuine competency-based tests (i.e., causally-related criterion measures) are largely independent of the content areas in which they are used. For example, the tests for thematic analysis, analysis of argument, problem solving, speed of learning, and other such measures described in the text test for generic abilities (competencies) which can be demonstrated in the context of any specific content area. These tests can be adapted to the natural sciences, social sciences, and humanities with equal facility; the content area does not determine the effectiveness of the test. We

---

[5]For example, vocabulary is correlated with college grades. However, one would not go about improving college grades merely by increasing vocabulary. Doing well in school requires abilities for problem solving, utilizing new information, and other skills not measured by vocabulary tests. Vocabulary is merely a tool, and how it is used depends upon other abilities and characteristics of the individual. One cannot do well in school without a reasonably adequate vocabulary, but having a strong vocabulary will not guarantee success in school without its effective use.

will always need tests of knowledge, but we also need tests of the way this knowledge is used. The measures discussed in the third section satisfy both of these criteria, which represent the essence of competency-based assessment.

A criticism leveled at the PBTE movement is that it is limited to preparation of "didactical technicians" (Broudy, 1972). A narrow correlational model of competence has fostered this notion, and this concern is legitimate to the extent that criterion validities depend exclusively upon specific task-oriented behavioral criterion references. Such validities for teacher education "are of sporadic interpretive utility" at best since they ignore the linking of test behavior to a more general attribute, process, or trait which provides an evidential basis for interpreting the *processes underlying test scores.* (Messick, 1975)

ICA strongly endorses this position but adds that construct validation is itself all too often limited in the types of referents it uses to provide meaning to test scores. Thus, we advocate a validation model that draws from the strengths of construct validation more heavily in the context of *real world events or life outcomes* than in the context of other constructs alone or "laboratory" behaviors. While Messick (1975) de-emphasizes criterion-referencing, he only does so (1) in terms of using criterion-referents outside of the context of construct validation and (2) perhaps in terms of the type of criterion used as referents. Indeed, all validation is criterion-referenced. The difference in criteria (e.g., "real world" performance, other tests, or observable "laboratory" behavior) determines the extent to which the meaning of the test responses are general or specific and of theoretical or real world significance. A difference between McClelland's (1973) and Messick's point of view is McClelland's emphasis on choosing real world behaviors as opposed to tests (which typically tap respondent rather than operant behaviors) and laboratory behaviors, as criterion referents. Thus, criterion-referents constituted by a nomological network of life outcomes are consistent with Messick's argument. Espousing such referents differs from Messick's point of view only in terms of emphasizing their selection as criteria for construct validation, not in the validation procedures or concepts themselves. In other words, Messick's notion of construct validation theoretically would include criterion behaviors, but empirically there are differences in emphasis on the types of behaviors to be included. It is for the sake of this difference in emphasis, not theoretical differences, that we have isolated real world events or life outcomes as critical factors in determining the real meaning of tests.

The strength and future of PBTE rests on its ability to support the rigorous type of research analysis which involves construct validation based heavily upon real world teaching outcomes. Until we have identified the critical intervening variables in the causal chain between the educational experience and performance in the field, we will be legitimately faulted by critics who view PBTE assessment (and education) as too narrow in scope.

APPENDIX C

THE PROBLEM OF ESTABLISHING CRITERION LEVELS OR PERFORMANCE STANDARDS

As the meaning of measures becomes established by construct-validation and empirical (criterion-referenced) links between education and the requirements of postacademic life, the question of what criterion levels of performance are necessary for granting credentials is made easier, perhaps, because concrete information exists with which educators can make sound judgments. Yet, the problem of establishing standards for levels of performance is a complex one because (1) this determination of appropriate levels of performance is dependent upon educators' goals for credentialing students, and (2) technical issues related to understanding the meaning of maximum levels of performance and the meaning of complex interaction of abilities probably necessitate highly subjective determinations of criterion standards.

With regard to the first point about determining standards of performance, Hodgkinson (1975) stresses the importance of asking good questions about the use and purposes of assessment. Sound judgment and planning are necessary to avoid proceeding with evaluative decisions based on ambiguous criteria, standards and/or levels of outcomes. These questions must include: *Who establishes criteria or standards*–an external auditing agency, a faculty member, the institution? *What is the reference group with which one will be compared*–performers in the real world, students in past years, other students currently being evaluated, one's own past performance, an *ideal* student? *What is the proper method of comparison?*–norm-referenced tests, criterion-referenced tests, behavioral measures, narratives (e.g., portfolios, diaries of past experience), unobtrusive measures, etc.? *What is the nature of the standard*–job performance in the "real world," individual growth and development, ideological ideals of performance, standardized scores? *What is the function of the standard*–to select or reject people, to improve performances, to admit students to professional schools or jobs?

If these questions are asked and the answers are concrete, specific and meaningful, a student should know who is judging him, how he will be judged, the nature of these judgments, the objectives related to them, and

how well he must perform to meet those objectives.

With regard to the second point about determining standards of performance, two conceptual or technical considerations are also relevant.

1. The Problem of Maximum Levels

Credentials are often restricted to those whose scholastic performance and/or test scores are higher than minimal levels required for work or other social roles. Such occurrences discriminate unfairly against those who are competent to work, for example, but who are selected out of occupational opportunities by those who believe in the simple equation: higher academic achievement means better work or life performance. The tacit assumption that *superior* abilities in *all* measured characteristics are necessary or even desirable for performance is highly questionable.[6]

Measures typically used to assess job task performance and performance relating to the mastery of units in a curriculum typically have little bearing on how subunits interact. For any given job, life task, or individual performance, component skills in one area can compensate for deficiencies in others creating a variety of combinations of individual performance levels which could theoretically add up to equivalent overall performance. Thus, minimal levels of performance on individual variables (which compromise overall competence) may have little meaning by themselves. Their interactions with respect to outcomes may have far greater significance.

We are most familiar with this problem in cognitive areas of education. We are often taught language use, verbal reasoning, spatial relationship, reading comprehension, abstract reasoning and syllogistic analysis (e.g., as measured by Miller Analogies) as discrete units of curricula. Assessment of integrated or general skills such as problem solving often do not take into account the interactive nature of skills in these subcomponent areas. Cognitive measures are used almost exclusively in assessment as if the qualities

---

[6] A simple motor skill example will demonstrate this point: we know that an automobile driver must grip the steering wheel with enough force to maintain control of the car. But beyond a certain level of pressure, added strength in holding the wheel does not increase overall driving competency. And this is just one of some 3,400 discrete behaviors identified by researchers as making up the task of "driving."

they measure did not interact, i.e., they are tested separately.

The importance of interactions, while intuitively obvious in the motor skills area, have not been carefully attended to in cognitive and social/emotional areas of assessment. Yet, once individuals have gone through a series of academic life experiences that enhance their competence in dealing with school, work, and other life experiences, the appropriate assessment task becomes that of measuring such integrated and generalized learning outcomes as the ability to cope with new problems, to find appropriate solutions, and to take correct actions.

Measures which reflect the interdependent nature of cognitive skills essential for satisfactory functioning outside of academia have only begun to be developed.[7] For example, Klemp's General Integrative Model of Assessment, incorporating a variety of independent techniques, is an approach to summative evaluation of an individual's ability to solve a problem which has as many elements and complexities of real life situations as possible. Such an assessment of individuals has the potential of coming closer to tapping real life competence than can any single test alone.

While it makes sense to require minimal levels of proficiency for many competencies, ability levels over and above necessary cut-off points do not always correlate with overall performance.

For example, in a job analysis, McClelland and Dailey (1974) found that a minimal level of organizational or clerical competence was necessary for human service workers in the Massachusetts Civil Service system, but *high scores on these measures were negatively correlated with superior job performance.* Selecting people by rank according to score not only discriminated against those whose scores were adequate (sufficient) though "Uncompetitive," but the process failed to select the better job performers as well. This finding and others[8] suggest that going beyond *sufficient* levels of com-

---

[7] A recent example in the noncognitive area by McClelland and Burnham (1976) reports the importance of the interaction between levels of motivation and ego-maturity for managerial competence.

[8] A recent study at Harvard revealed that the past SAT scores of faculty members were *negatively* correlated with more successful teachers (McClelland, personal communication).

petency in awarding credentials can be very dysfunctional for society–not only in terms of equity, but in terms of meritocracy as well.

In many job situations, where cognitive and other competency measures are used to select job applicants, even if job relevance of the characteristics being tested for can be demonstrated (e.g., "verbal ability" in human service workers), *level of sufficiency* for competent job performance is rarely evaluated or known.

We need more empirical research to establish *minimal levels of competence required* for quality performance *based on how workers in the field perform* on various competency measures.

2. The Problem of Interactions

Researchers have long recognized that the interaction effects of variables are quite often more significant and meaningful than individual variables taken alone. It was stressed earlier that competence is not a simple summation of discretely defined skills and abilities. This is readily seen in the example of driving ability. Although one can identify many skills necessary for safe and effective driving–including attitudes, cognitive skills, and emotional factors, as well as perceptual and motor skills--it is intuitively obvious that a simple summation of measurement scores on these discrete task performances would not add up to equivalent driving skills. An individual who is overly competent at some driving skills but woefully inadequate at others would be a poorer driver than someone whose skills were all sufficient, though their summed skill scores would be identical.

The implication for higher education is that one cannot assume that abilities or skills discretely learned will be integrated in work and life functions and consequently that establishment of minimal levels of performance on isolated skills or "subcompetencies" have much meaning in themselves. Therefore, competency research, new assessment procedures, and test instruments must also focus on the interdependence of skills.

APPENDIX D

REFERENCES

American Association of Colleges for Teacher Education, Committee on Performance-Based Teacher Education, *Achieving the potential of performance-based teacher education: Recommendations,* PBTE Series: no. 16. Washington, D.C.: American Association of Colleges for Teacher Education, 1974.

Atkinson, J.W. *Motives in fantasy, action, and society.* New York: Van Nostrand, 1958.

Broudy, Harry S. *A critique of performance-based teacher education,* PBTE Series: no. 4. Washington, D.C.: American Association of Colleges for Teacher Education, 1972.

Broudy, Harry S. Can research escape the dogma of behavioral objectives? *School review,* 1970, *79,* 43-56.

Dodl, Norman et al. Florida catalog of teacher competencies. Albany, N.Y.: Multi-State Consortium on Performance-Based Teacher Education, 1973.

Gage, N. Models for research on teaching, Occasional paper no. 9. Stanford: Stanford Center for Research and Development in Teaching, School of Education, 1976.

Hodgkinson, H.L. Evaluation to improve performance. In Vermile, D.W. (Ed.), Learner-center reform. *Current issues* in higher education 1975. San Francisco: Jossey-Bass, 1975.

Klemp, G. *The objective measurement of prejudice and discrimination: Final report to the Office of Naval Research.* Boston: McBer & Co., 1975.

Klemp, G. *Validation of a broad-based examination of entry-level support positions in administrative and technical work.* Boston: McBer & Co., 1976.

Kohlberg, L. Stages of moral development as a basis for moral education. In Beck, C. & Sullivan, E. (Eds.), *Moral education.* Toronto: University of Toronto Press, 1970.

Kolb, D.A. On management and the learning process. *Working paper no. 652.* M.I.T., 1973.

Kolb, D.A. Individual learning styles and the learning process. *Working paper no. 535-71.* M.I.T., 1971.

Kolb, D.A. and Goldman, M. Toward a typology of learning styles and learning environments. *Working paper no. 688-73.* M.I.T., 1973.

McClelland, D.C. (Eds.) *The impact of college on personal characteristics and life outcomes.* Cambridge, Ma.: Harvard University, 1976 in press.

McClelland, D.C. *Power: The inner experience.* New York: Irvington Press, 1975.

McClelland, D.C. Testing for competence rather than for "intelligence". *American psychologist, 28,* 1973.

McClelland, D.C. Achievement motivation can be developed. *Harvard business review,* 1965 November-December.

McClelland, D.C. *The achieving society.* New York: Van Nostrand, 1961.

McClelland, D.C., Baldwin, A.L., Bronfenbrenner, U., and Strodtbeck, F.L. *Talent and society.* New York: Van Nostrand, 1958.

McClelland, D.C. and Burnham, D.H. Power is the great motivator. *Harvard business review,* 1976 March-April.

McClelland, D.C. and Dailey, C.A. *Professional competencies of human service workers: A report to the College for Human Services.* Boston: McBer & Co., 1974.

McClelland, D.C. and Klemp, G. *Report to the Executive Office of Manpower Affairs on the validation of a human service worker test.* Boston: McBer & Co., 1974.

McClelland, D.C. and Winter, D.G. *Motivating economic achievement.* New York: The Free Press, 1971.

Medley, Donald M., Soar, Robert S. and Ruth *Assessment and research in teacher education: Focus on PBTE, PBTE series no. 17.* Washington, D.C.: American Association of Colleges for Teacher Education, 1975.

Merrow, John G. *The politics of competence: A review of competency-based teacher education,* a report to the National Institute of Education. Washington, D.C.: Institute for Educational Leadership, 1974.

Merwin, J. *Performance-based teacher education: Some measurement and decision-making considerations, PBTE series no. 12.* Washington, D.C.: Association of Colleges for Teacher Education, 1973.

Messick, S. The standard problem: Meaning and values in measurement and evaluation. *American psychologist,* 1975, *30,* 955-66.

Pottinger, P. and Klemp, G. *Final report to FIPSE, contract no. OEC-0-74-9258.* Boston: McBer & Co., 1976.

Pottinger, P. *Comments and guidelines for research in competency identification, definition and measurement, no. SURC-TR75-582.* Educational Policy Research Center, Syracuse University, 1975.

Rosenthal, R., Archer, D., Koivumaki, J.H., DiMatteo, M.R. and Rogers, P. Assessing sensitivity to nonverbal communication: The PONS test. *Division 8 newsletter,* American Psychological Association, January 1974.

Stewart, A.J. Longitudinal prediction from personality to life outcomes in college educated women, unpublished doctoral dissertation, Harvard University, 1975.

Stewart, A.J. *Scoring reactions to rhetoric of analysis of argument.* Unpublished manuscript, Harvard University, Department of Psychology and Social Relations, 1974.

Stewart, A.J. *Scoring systems for stages of psychological development.*

Unpublished manuscript, Harvard University, Department of Psychology and Social Relations, 1973.

Stewart, A.J. and Winter, D.G. Self-definition and social definition in women. *Journal of personality*, 1974, *42*, 238-59.

White, R.W. Motivation reconsidered: The concept of competence. *Psychological review*, 1959, *66*, 297-333.

Winter, D.G. *Scoring the test of thematic analysis for critical thinking.* Unpublished manuscript, Wesleyan University, 1973.

STATE INTEREST AND

TEACHER EDUCATION PROGRAM APPROVAL

By Lawrence D. Freeman

Larry Freeman completed his Ph.D. in English in 1968 at the University of Nebraska-Lincoln. During his doctoral studies he worked at the Nebraska Curriculum Development Center on the "Nebraska" Curriculum in English and assisted in establishing and implementing the Tri-University Project. After teaching English at SUNY at Binghamton for two years, Freeman served as Co-director of the Nebraska Confederation for Early Childhood and then as Associate Director of the SCUEET. He is presently Director of Teacher Education Program Approval in the Illinois Office of Education.

During his tenure at the Study Commission, Freeman focused much of his time and energy on examining the ways in which existing legal and quasi-legal structures may impede the development of better teacher education. In his present position, he continues his exploration in these areas. He is presently preparing a manuscript that extends the observations and arguments that he makes in the following essays.

The following papers are working papers presented to the State Teacher Certification Board of Illinois; they represent "staff thoughts" and do not necessarily reflect Illinois policy.

GENERAL DESCRIPTION OF THE ACTIVITIES

ASSOCIATED WITH REVIEWS OF

INSTITUTIONS AND PROGRAMS

* * * * * * * * *

CHARACTER AND USE OF EVIDENCE

IN PROGRAM APPROVAL

* * * * * * * * *

DUE PROCESS FOR INSTITUTIONS AND STUDENTS:

IMPLICATIONS FOR PERIODIC REVIEWS

STATE INTEREST AND

TEACHER EDUCATION PROGRAM APPROVAL

By Lawrence D. Freeman

At present, those administering state statutes regarding the certification of school personnel find themselves forced, or electing, to rely on highly generalized statements about an individual, on what I will call a "summary account." States administer statutes that typically call for a stated degree comprising a certain number of semester hours, awarded by a certain kind of institution; in some cases, the distribution of some of the hours is specified in a general way. The summary accounts of an individual's preparation and past performance show up as transcripts. In the typical case of a state statute regarding certification, two problems emerge: (1) credit hours range from meaningless to unreliable, and (2) it is increasingly difficult to sensibly distinguish between "professional" and "general" education, between "clinical experiences" and "student teaching." As a result, state officials—while trying to pursue actions borne of common sense—almost inevitably end up administering a set of stipulative definitions borne of administrative necessity that have an increasingly questionable relationship to performance on the job. Consequently, state officials typically become like high priests who alone have access to the arcane mysteries of certification.

What appears to be missing in the typical state statute, and in the rules and regulations, is an explicit notion of what distinguishes a teacher from a non-teacher. Or put another way, there is no easily accessible explanation of why the state has sought to exercise its police power requiring certification in order to be employed as a teacher in the public schools. A partial explanation of this phenomenon may lie in the peculiar validity of the teaching license. Teaching is a spontaneous human activity whose origins can be traced into pre-history. Only since the middle of the nineteenth century have claims for

---

Note: The views and opinions presented in this paper are those of the author. No inference should be made that they represent, in whole or in part, the official or unofficial views of the Illinois State Board of Education, the Illinois Office of Education, or the Illinois State Teacher Certification Board.

teaching as a "science," an art, or complicated set of skills been made. But the fact remains that the validity of a teaching license is restricted to a particular physical situation—the public schools. Outside of such situations, one can proclaim himself a "teacher" or engage in teaching without legal penalty. Unlike some professions and trades, the teaching profession has no enforceable definition of work and unlike, say, psychologists, the use of the title "teacher" needs no licensure. Thus, because the validity of licenses awarded school personnel is restricted to a physical setting and to eligibility to be placed on a payroll, the tough questions regarding the characteristics of "teachers" as opposed to "non-teachers" have typically not been addressed in the creation and administration of the teacher certification statutes.

We find ourselves, then, making distinctions between "teachers" and non-teachers on the basis of summary accounts such as transcripts. These summary accounts find meaning as a result of a complex web of stipulative definitions that in the view of some approach arbitrariness and irrationality and which under the day-to-day pressures of issuing certificates become increasingly removed from the naturally articulated concerns of parents, citizens, and the teaching profession itself.

Reliance on summary accounts will in all likelihood persist, if only for bureaucratic reasons. A system relying on more extensive, more detailed, and consequently more meaningful accounts is necessarily more unwieldy, more vulnerable to attack, and more expensive. Such a system probably requires a collective decision-making process because virtually every case represents a novel situation, something that can be fairly easily avoided when relying on highly generalized accounts.

We have come to rely on the credit hour, even though such reliance has been questioned almost since the use of credit hours became a widespread practice. The Carnegie Foundation for the Advancement of Teaching, a prime mover in developing and implementing the credit hour, began in the thirties to seek a replacement and expended considerable effort and money to develop the Graduate Record Examination. This whole effort might well be regarded as replacing one kind of summary account with another. In teacher education, an analogous effort resulted in the National Teachers Examination which purports to provide a more detailed accounting of an individual's achievements than a transcript. Similar claims are apparently made for those examinations used by jurisdictions such as Chicago and New York City.

Others at this conference will undoubtedly consider the validity of such claims and the relevance of these claims to awarding certificates.

In summary, it appears that those who have made and continue to make decisions in the process of awarding certificates have sought strenuously to develop ways to make the use of general and summary accounts reliable bases for their decisions. In spite of these efforts, however, these accounts remain based on examination of bits and pieces of behavior, not on an examination of a totality of behavior or entire acts. And, at the end, one does not know to what the summary account attests. It may mean only a cooperative registrar.

There are issues that reach beyond the adequacy of summary accounts. For instance, whatever kind of accounts one wishes to rely on, there is the issue of the relevance of the account to the decision to be made. This, in short, is the question raised by equal employment legislation, *Griggs* and its progeny: Do the summary accounts used as a basis for certificating meet any reasonable standards for admissability as evidence, and if they do, what kinds of inferences can legitimately be drawn from them, particularly inferences of a predictive character?

What we may need is an idealization of the certification system. From my perspective, such an ideal system would minimally meet the following criteria:

(1) The information and evidence considered would be sufficiently comprehensive and detailed that they provide a picture of the totality of an individual's competence and ability. Well-designed decision-making in regard to tenure might illustrate what I have in mind here.
(2) The information is admitted for consideration on the basis of complying with a set of standards rationally related to the class of jobs the certificate makes one eligible for.
(3) Inferences drawn from the information are rational, legitimate, and logical. Such inferences need not be validated by empirical evidence but have at least an aura of common sense.

Transcript analysis fails on all three levels; exclusive or partial reliance on paper and pencil tests also fails. What then is to be done?

At the present moment, a mandated system of institutional and program approval seems the best way to proceed. Personally, I believe program approval and issuing certificates by entitlement may be legally vulnerable. I remain skeptical of the legal soundness of delegating certification responsibilities to public and private colleges and universities and of other features of the arrangement. But faced with the trade-offs, I am convinced that program approval represents the only process that minimizes the legal vulnerabilities and simultaneously has the potential for approximating the criteria I set forth above. Program approval, for instance, puts meaning behind transcripts and degrees and fixes the responsibility for assessing a candidate's total behavior at a level where such assessments can be made.

Program approval, however, is not presently without its liabilities and vulnerabilities. Some of its liabilities can be traced to its historical reliance on the kinds of standards and technologies developed in the sphere of voluntary accreditation. Voluntary accreditation has always had a different set, or at least additional, objectives than state-mandated systems of approval. But program approval can also learn from recent developments in voluntary accreditation. For instance, the development of appeals processes in voluntary accreditation might be pursued. Such processes should, with care for preserving due process, enable challenges by institutions, candidates for certificates, prospective employees, and consumers.

Program approval, to realize its potential, must also be prevented from becoming merely transcript evaluation writ large. It must avoid the rigidities of such evaluation in order to facilitate institutions' quick response to changing and emerging needs in the public schools. And of course program approval must be supported with sufficient resources that rigorous monitoring can occur.

But most importantly, program approval systems, in order to do their job, need to become *public* and *adversarial* in character, and they need to be based on some clear conception of the *state interest* that they purportedly protect. It is perhaps the lack of a conception of state interests to be protected that renders present program approval systems suspect; and one might add that the recent overcoming of the chronic shortage of teachers during the past few decades and the sudden creation of an oversupply have effectively called into question the relationship between program approval and conceptions of state interest. As long as there was a chronic shortage of teachers, it

was at least arguably in the state interest to approve institutions and programs known to be demonstrably deficient even in basic ways. But the present situation may mean that other expressions of state interest can and will assert themselves. And this leads to the question whether there currently exist or can exist systems of program approval under which institutions and programs are denied approval on grounds of not fulfilling state interests.

I wish to come at this question indirectly, first creating a context which will hopefully move the question from its pure abstract form to the realities of day-to-day program approval. As I have noted, program approval shares many characteristics with private accrediting. One of the most deleterious features shared by the participants in both processes is the tendency to conduct business in private and avoid public disclosure. Since both processes are in fact, if not in theory, engaged in public business or business strongly affecting the public, the processes and actions emerging from them must see the light of day. Mitchell Wendell has observed:

> Teacher certification agencies operate processes in which the officials and employees may in fulfillment of their human nature prefer not to respond to inquiries about the types of records they keep or their content. This is likely to be in relation to the criteria they employ in determining whether particular applicants should or should not receive certificates and the weight they assigned to pieces of information or allegations contained in individual files. It is probably not that many of the decisions would likely be found wrong. Rather, it is that they might be troublesome to document or justify in a way that would appear to all as scientific and objective. ("Freedom of Information Statutes, Teacher Certification Agencies, and School Systems," State Education Department, Interstate Certification Project, New York)

What applies to individual certification cases may apply as well to institutions and programs. I have no evidence that program approval is not public in all states. I can only testify that I have personally been subjected to intense pressures from a variety of sources to keep out of the public record items tending to embarrass institutions and others in a system that one might characterize as only minimally public.

I know of no system that assures that program approval staffs will

resist the temptation to adopt a policy of consensus and appeasement with regard to institutions. Integrity and courage certainly assist. But perhaps the creation of public bodies that review all the available relevant information and the judgments of staff persons and come to their independent conclusions in public will augment integrity and courage and assist in replacing appeasement with a more defensible strategy.

In addition to creating program approval systems open to meaningful public inspection, program approval systems must in my view become more *adversarial* in character if they are to realize their potential. Given the present character of many program approval systems, it is virtually impossible to pretend that the record concerning an institution and its programs is complete, accurate, and valid. Typically, the sources of information are limited to the institution itself, visitation teams, and state agency files. The visitation process, a key component in this whole process, is unavoidably prone, it seems, to producing tons of irrelevant or virtually meaningless prose and to bias, and is vulnerable even to outright misrepresentation. At present, the program approval process depends almost exclusively on human testimony and consequently has all of the problems attendant to relying on human testimony when making decisions. But typically program approval systems have none of the protective mechanisms available in other areas where human testimony is heavily relied upon.

Now, let's be clear that human testimony must be relied upon. Any meaningful evaluation of any kind of education program must, it appears, rely at least to some extent on human testimony. The problems raised by human testimony need not lead to excluding it from consideration but rather to insuring that it is legitimately employed in balancing the interests and rights of institutions and others against state interests.

I have taken this long way around the question I raised earlier—on what basis does one deny approval to an institution or program—in order to set in a meaningful context some more dramatic issues. Before issues relating to the use of human testimony arise, there are some other problems needing examination. They can be formulated easily but resolved with only the greatest of difficulty: (1) What issues are involved in approving or disapproving an institution? and (2) What are the points of controversy? Theoretically these problems should be easily answered because one would assume that reference to the state interests served by certification would define the issues and

points of controversy. It should be a simple task to say what is at stake in the approval process, if only by reviewing legislative statements as to what schools are encouraged or mandated to do. Inspection of relevant statutes and rules and regulations is, however, frustrating. Not only do patent conflicts appear, but state legislatures are prone to mandate teaching of the contributions of ethnic groups, conservation, state history, citizenship and patriotism, bilingual education and so on. These expressions of state interest frequently do not coincide with the departmental structures of higher education or paradigms of twentieth century knowledge. And judicial expressions of state interest have historically tended to stress citizenship and preparation for effective participation in social, political and economic processes. One opinion, for instance, argues that:

> We are a self-governing people, and an education prepares the boys and girls for the duties and obligations of citizenship. Neither the schools nor the state can carry on without rules or laws regulating the conduct of the student or citizen, and those who are taught obedience to the rules and regulations of the school will be less apt to violate the laws of the state.

At best, the relationship between program approval and putative state interests is not easily articulated; at worst, one cannot even discern it. We can perhaps best recognize this state of affairs and settle for a less than ideal and vaguely stated policy that what is at stake in the approval process is whether the institutions and programs are responsive to public school staffing needs. If even this vague policy statement is adopted, the issues and points of controversy become clearer. But of course the controversy is not settled because one must rely on human testimony to make the necessary determinations.

The problems created by human testimony range from competence to credibility, from bias to the failure to provide, or deliberate withholding, of evidence of acceptable quality. Certainly one has to recognize that human testimony gains its worth and credibility from the human being providing it.

And these facts create additional problems because at present in the systems of program approval I am familiar with, assessments of the "human being offering" the testimony are virtually impossible for institutions, decision-makers, and others who have interests or responsibilities. There is little possibility for an institution, for instance, to challenge an evaluator's

perceptions and for decision-makers to tease out the nuances of his statements. While a system of cross examination appears impossible, at least at present, I suspect we can improve the state of things by enabling institutions and other interested parties—individual teachers and their organizations, parents, prospective employees, students—to respond in writing to the evaluator's perceptions and statements, include those responses in the record, and seriously consider them in the review process.

While this creates a somewhat messy procedure—conflicts and embarrassment, and even bad manners may frequently occur—it appears to me that if these conflicts can be brought home on questions having to do with state interest in the certification process, the approval system itself seems to clarify the state interest or at least outline competing theories of state interest.

Perhaps the most dramatic but increasingly real way of focusing the questions we are dealing with is consideration of this question: What is the character of the evidence necessary for an official or public body to declare that an institution and its programs are not serving the state interest and therefore ought to be denied the right to train teachers? Or, one can put the question in even a more troublesome way: What is the character of the evidence necessary to declare that state interests are being served?

We have, in the past and perhaps continue, answering these questions by relying more on rules than on what I will refer to as standards. We have tended to set out prespecified behavioral statements, on the order of rules governing driving automobiles. We have assumed that rules analogous to those requiring drivers to stop at red lights have protected the general welfare and promoted state interests. Such rules are nice because one need not make any inferential leaps, for in all cases the "facts speak for themselves." And in talking to obstreperous clients on the phone, one can be at ease knowing full well that the rule and evidence for compliance are synonymous. The bureaucrat on the phone does not have to engage his mind defending any inferences and those of us who have received the same calls can continue to ignore the fact that in considering the protection of state interest all one has at hand is circumstantial evidence. We can thus deny that somewhere along the line we depend on suggestion, analogy, inference, metaphors, and, yes, even subliminal cues.

Rather than relying on narrowly construed rules, I would argue that we

seek to enforce standards rather than rules. While more ambiguous, standards at least admit of the possibility of raising significant questions, of engaging the judgments of several individuals, and actively seeking to render certification as a means of protecting state interest, which is probably dynamic rather than static. Rather than specifying in advance what evidence will be persuasive, we need to be open to an adversarial context. This, of course, means that institutions will have to develop the guts to stand up and fight for programs they believe serve the public interest, and one can foresee professional associations, community groups and others doing the same. And, such a strategy of course means that state departmental staffs must be made and held accountable to some sort of public body and forced to reply to positions advocated by others. Such a context would quickly produce some hard work on issues that continue to be ignored—what kinds of evidence can be allowed to be brought under consideration and what kinds of inferences can be made in attempting to determine whether an institution and its programs are serving the state interest.

I am aware of the temptation to close a paper like this one with a plea for further research. But I suspect that the solutions to the difficulties we face are more likely to arise in less disciplined and more fascinating ways. I would suggest that in working on these problems we do the following:

(1) attend to the naturally articulated, if frequently imprecisely stated, concerns of teachers, parents, administrators and others;
(2) assume that "state interest" is not a static thing with a logical character but dynamic and socially and historically produced;
(3) that the "education of educational personnel" embraces a wide variety of processes and is not a single process for which "reality" can be claimed; and
(4) that continuous critical reflection on our day-to-day efforts—our ordinary talk and behavior—may provide direction in devising standards and systems enabling us to make more reasonable and useful decisions.

I wish to acknowledge the influence and assistance provided by Robert L. Wolf's provocative dissertation: "The Application of Select Legal Concepts to Educational Evaluation," University of Illinois, December, 1973.

GENERAL DESCRIPTION OF THE ACTIVITIES ASSOCIATED WITH REVIEWS OF INSTITUTIONS AND PROGRAMS

By Lawrence D. Freeman

Tenth Year Periodic Review

The Tenth Year Periodic Review falls into four phases: (1) preparation of the institutional report, (2) the team visitation and validation of its report, (3) the assessment of available information and development of recommendations by the staff, and (4) action by the State Teacher Certification Board and Superintendent.

Preparation of the Institutional Report

The Institutional Report should be a succinct and accurate description of the institution and its teacher education programs. It should be sufficiently detailed to provide adequate information for assessing whether the institution and its programs comply with the standards and criteria. The standards address mainly the place of teacher education within the institution and its commitment to teacher education; the criteria are more specifically focused on programs. Program as used here refers to a structured sequence of learning activities and experiences designed to lead to a specific certificate and endorsement. Clear distinctions between required and elective components in each program should be made including academic and in clinical work. A description of the integration of general education, academic specialization and professional education for each program should be presented.

The report should cover unique aspects of the institution and any external relationship such as contractual relationships with other IHE's or with LEA's for teacher education purposes and any governing Board policies, e.g., the Board of Higher Education mandated relationship to community colleges. The report should also make clear distinctions between "what-has-been-and-is-being-phased-out," "what-is," and "what-is-planned." And the

report should indicate throughout, if possible, where written documentation for various aspects of institution and programs can be found.

There is no imposed format for the report. Difficulties tend to arise in four areas: fiscal data, faculty data, library and curriculum materials, and ancillary and off-campus activities. In all cases, the data presented should be determined by the relevant standard or criterion. Summary of data with trends or general characteristics should be sufficient, with more detailed information available to the visitation team as needed and requested. Ancillary and off-campus activities reported should be presented as part of, or otherwise directly related to, programs being reviewed.

During the preparation of this report—which frequently begins one and a half years prior to the date of the visitation—IOE staff are assigned to work with the institution. The character of this liaison work is that of technical assistance. The staff meets with institutional representatives to explain the meaning of the standards and of the criteria and the kinds of information necessary to make the determinations that will have to be made during the program review process. In most instances, such liaison requires actual presence on a campus twice, sometimes more frequently, and meetings between institutional representatives and staff in Springfield. It is during these liaison activities that the staff begins to achieve an understanding of any potential difficulties or excellencies that the institution may possess. Records of liaison contacts are maintained so that there is a complete record of what has transpired between the staff and the institution. It should be noted that the staff in no way writes the report for the institution or does more than offer assistance in suggesting kinds of information or evidence and potential sources for the information and evidence that should be presented in the report. We regard this phase of the Tenth Year Review as crucial because it begins the development of a record on which decisions will be made. The staff is concerned that the institution be accurate and represent the state of affairs as it actually is at the institution at that time.

Visitation

The visitation team is made up, in accordance with State Teacher Certification Board rules and regulations, of school teachers and administrators, and college and university faculty and administrators. The selection of team members is done by an IOE staff member who is designated by the

Director as the chairperson. The team members are approved by the Director. All team members are from Illinois, and where possible (in accordance with State Teacher Certification Board policy) include a State Teacher Certification Board member. In the case of public institutions, a staff member from the governing Board may also be asked to serve on the team in an *ex officio* capacity. In all cases but those of the State Teacher Certification Board member, the chairperson and any other IOE staff and staff from a governing Board, the institution and each team member is asked to indicate whether there are past or present relationships between the institution and members of the team which constitute, or may be construed as, a conflict of interest. When grounds for a conflict of interest are alleged they must be set forth in writing. In no case can an institution veto team members. IOE assumes responsibility for assembling a competent team, though it does not pretend to select "unbiased" team members. The team will have enough members to examine each and every program in detail. No team member will review more than two programs in normal circumstances. The team is instructed in writing by means of a general charge to the team, and each team member receives a written charge regarding his specific responsibility. These charges are shared with the institution. These instructions indicate that the visitation team and its individual members are to engage in a fact-finding effort designed to generate information sufficient to determine whether an institution and its programs comply with the standards and criteria.

The team is not, as is the case in voluntary accreditation, a recommending body and does not determine whether the institution is in compliance. The team is directed to investigate relevant written documentation and to interview students, faculty, cooperating school personnel, administrators, and other persons deemed to have relevant information concerning the institution and its programs. After each team member has filed a report with the chairperson, the IOE staff lightly edits materials supplied—deleting only clearly irrelevant statements and corrects grammar and spelling. It is then sent to team members for corrections, additions, or deletions. It is simultaneously sent to the institution. The institution is asked to make written responses to errors in fact indicating where and how its version of the facts can be substantiated. Where the institution feels that bias or a perspective alien to its own has been imposed by a team member, the institution is invited to respond setting forth its view of the matter. Where the institution's version of the facts is substantiated, corrections are made in the report. In other cases, the institutional response is incorporated into the team report and noted as

an institutional response. Institutions are given one month to six weeks to respond.

Staff Analysis and Recommendations

The purpose of the team report, as well as the institutional report, is to prepare a record of information that is comprehensive, accurate, complete, and that oftentimes includes conflicting points of view. The staff reviews the institutional report, the team report, and any other relevant information that has come to it in terms of the standards and criteria, State Teacher Certification Board rules and regulations, and relevant statutes. The staff prepares an analysis for the State Teacher Certification Board and the Superintendent in which it discusses the information in terms of the standards and criteria and in which it sets forth recommended action by the State Teacher Certification Board and Superintendent. Recommendations are made concerning the institutions, recognition for teacher education purposes and for approval of each program for teacher education. During this process of preparing the staff analysis and recommendations, staff (where uncertain or where conflicts arise) is in touch with both representatives of the institution and individual team members as it seeks to work out difficulties. Staff analyses are drafted by a staff member who has not served in a liaison capacity with the institution, has not chaired the team, or participated on the team. This draft is then reviewed by the Director, often rewritten, and frequently given to another staff member to challenge, with yet another draft prepared. After the results of this process are in, oftentimes all those involved get together to work out any differences. The usual concern is the relationship of evidence to recommendations and construing the standards.

Action by State Teacher Certification Board and Superintendent

The Superintendent, in consultation with the State Teacher Certification Board accepts, rejects, or modifies the staff recommendations. These actions, including any appeals by the institution, are described in the *Manual* adopted in March, 1975.

Fifth Year Periodic Reviews

The Fifth Year Periodic Review falls into three phases: (1) the preparation of an institutional report, (2) staff review and preparation of an analysis

and recommendations, and (3) action by the Superintendent in consultation with the State Teacher Certification Board.

Preparation of the Institutional Report

The preparation of the institutional report proceeds in the same way that reports are prepared for Tenth Year Reviews. In some cases, however, NCATE reports are used in lieu of a separate report prepared especially for the Fifth Year Review process. In non-NCATE cases, the staff provides liaison and visits a campus usually at least twice during the preparation of that report. In NCATE cases, the staff will serve on the NCATE team and frequently is involved in liaison activities as the institution prepares its NCATE report. If the NCATE report does not, in the judgment of the staff, provide sufficient information for the Fifth Year Periodic Review, the institution is requested to prepare an addendum. In some cases the staff also has access to the NCATE team report and to materials that an institution may have prepared in meeting with representatives of NCATE.

Preparation of Staff Analysis and Recommendations

This process is much like that used in the preparation of staff analysis and recommendations for the Tenth Year Review process. The staff review focuses in three areas: (1) compliance with the requirements for certification imposed during transcript evaluation, (2) whether the institution and its programs comply with the standards and criteria as can be determined in the absence of a visitation team report, and (3) contacts the institution requesting further information clarifying items about which the staff is unsure or unclear.

Action by the State Teacher Certification Board and Superintendent

In accordance with the *Manual*, the Superintendent (in consultation with the State Teacher Certification Board) acts to accept the Fifth Year Report. If the report is unacceptable, the Superintendent (in consultation with the State Teacher Certification Board) may withdraw recognition of the institution or approval of one or more of its programs.

New and Modified Program Reviews

In the case of proposals to sponsor a new or modified program, the

institution first notifies the Secretary of the State Teacher Certification Board that the institution intends to develop a new program or to modify an already existing one. At this point a staff member is designated to provide liaison to the institution. The liaison is of a technical assistance character with the main activities being explaining and explicating the standards and criteria and the processes through which the proposal must go. This liaison involves visitation to the campus and a general overview of resources and interviews with faculty and administrators. The concern of the staff here is to insure that the institution presents sufficient and accurate information so that it can make a recommendation. Under the procedures outlined in the *Manual*, the institution's designated officer must request of the Secretary of the State Teacher Certification Board that its program be acted upon by the State Teacher Certification Board and the Superintendent. The preparation of staff analysis in the case of new programs involves a review of the program against the standards and criteria set forth in the *Manual*, state statutes, rules and regulations, and policies that have evolved through the interaction of the Superintendent, the State Teacher Certification Board, and the staff.

CHARACTER AND USE OF EVIDENCE IN PROGRAM APPROVAL

By Lawrence D. Freeman

In review of institutions and programs, a number of issues surrounding the availability and use of evidence come to the fore. Four such issues are illustrated below: (1) Kinds of Evidence, (2) Reliability of Evidence, (3) Relevance of Evidence, and (4) Sufficiency of Evidence.

I. Kinds of Evidence

An examination of institutional and visitation team reports reveals that most evidentiary material can be classified as one of four kinds:

A. Verifiable factual evidence is usually of a quantative or documentary nature, consisting of transcripts, records, minutes, and so forth. This evidence qualifies in most instances as direct evidence in support of claims.

B. Human Testimony
   1. Direct human testimony provides a description of the report writer's observations, perceptions, and activities.
   2. Inferences usually appear in the report when a writer presents judgments about a program or one of its aspects, or when the writer tries to resolve conflicts in evidentiary material he has gathered.
   3. Indirect human testimony consists of the writer's reports and summaries of interviews and conversations. Examples:
      a. Advisors in each department are given a form to complete concerning their assessment of the candidate's potential for success in a teacher education program. (Factual Evidence)
      b. It was reported that no student is admitted to the teaching program without the approval and written assessment of the history department. (Indirect Human Testimony)
      c. An examination of some student files in the education office confirmed that some students from the history de-

partment are admitted to the teacher education program with the form incomplete. (Factual Evidence)

d. The bottom line of this form contains room for written comments. Those observed were short and superficial. (Direct Human Testimony)

e. The possibility exists that the department of history may not give its approval to the student. If this occurs, it may mean that the assessment by that department is viewed as more reliable than the performance data compiled by the department of education. (Inference)

II. Reliability of Evidence

Several issues repeatedly occur in assessing the reliability of various kinds of evidence. Evidentiary statements frequently do not occur in "pure" form as in the examples above, and even if they do, the sources for evidentiary material become crucial factors. After reviewing a number of team reports, the staff has selected the following statements to illustrate these difficulties.

A. Undesignated source with no elaboration: "There is evidence to suggest that the general education program serves students well." Commentary: By itself this sentence is virtually meaningless. The reader of this sentence has no clue as to the kind of evidence referred to or the sources of evidence. Such highly generalized statements are regarded by the staff as highly unreliable.

B. Summary of a number of conversations and interviews with designation of source by title: "Interviews with students, graduates, teachers, and school administrators revealed that most students are inadequately prepared for student teaching." Commentary: This statement is stronger than the preceding one, but it raises a number of questions, such as how many persons in the various roles were spoken to, how did the writer elicit this information, and was the testimony as consistent and unanimous as the statement implies?

C. Designated Source: "Dr. Smith indicated that he has been in constant contact with local school personnel. Three teachers, Ms. Barrow at Willow, Mr. Clark at King, and Ms. Ogden at Middlebrook, confirmed this." Commentary: Such a statement inspires a

high degree of credibility. The statement is specific with respect both as to what is being asserted and sources of information.

D. Inference: "This is an excellent procedure, but not if it is to the exclusion of additional experiences with the age level of children in which they are most interested." Commentary: This statement clearly offers the writer's judgments on an aspect of the program. Its reliability will probably be governed by the context in which it appears; that context is likely to indicate whether the writer is given to unwarranted bias or is a reliable and perceptive observer. Oftentimes inferences made by writers appear to be insignificant upon first reading but come to be extremely important since they serve to assign relative importance to other kinds of evidence.

E. Team Member's Testimony: "Dr. Smith's claim that classes are very small and the instructor is able to do much informal evaluation of students was substantiated by visits to seven classes." Commentary: This statement is regarded as highly reliable. One has the sense that the writer is tapping a number of sources of information and is seeking to develop an accurate and complete record.

F. Documentary Evidence: "The catalog states that students in this program are required to complete a course in the methodology of teaching social studies, a fact confirmed by examination of transcripts." Commentary: In this instance the statement appears to be incontrovertible. Two sources of documentary evidence provide consistent information.

The commentary about each of these examples ignores the contexts in which they appear. Frequently, statements of apparently low reliability gain credence when essentially the same observation appears with differing degrees of reliability throughout the team report.

Assessing the reliability of evidence is difficult, particularly when confronted by conflicts in the evidence. This is particularly the case when, for instance, commentary by students and cooperating teachers, who are often unidentified, conflicts with data gathered by a program evaluation system, with records maintained by the institution, or even with commentary by identified representatives of the institution. Reports concerning cooperating teacher and student perceptions are particularly difficult to use at face value because it is often unclear how many people were talked to and whether their statements were virtually unanimous.

A constant concern of the staff is how to resolve conflicts fairly and how to determine what kind of evidentiary material should be given preference.

III. Relevance of Evidence

In some instances, the relevance of evidence to one or more standard or criteria is not difficult to establish. This is particularly true when treating the standards and criteria as "rules." There either are or are not objectives, methods courses, required clinical experiences. In other instances, however, the relevance of evidence is not so clear cut. For instance, the criterion that requires "that the program provides for acquisition of knowledge . . . necessary for effective performance in specific teaching . . . roles," frequently creates difficulty. Usually two issues arise: the character of the role for which the person is being prepared and what is needed to perform effectively in it. For instance, a team member observes that an English program does not require instruction in reading and that secondary English teachers frequently encounter reading difficulties in the classroom. In responding to such observations, the staff does the following:

A. Insures that statutory requirements and State Teacher Certification Board rules and regulations are met.
B. Insures that A-160 requirements and reimbursement requirements, if any, are met.
C. Tries to determine if the institution is evaluating its program and whether that data suggests requiring a reading course.
D. Consults with Illinois Office of Education staff who work with school personnel in the areas of English and reading.
E. Reviews available professional literature.
F. Reviews past actions in similar situations.

IV. Sufficiency of Evidence

Determining whether or not there is sufficient evidence to sustain a recommendation is an extremely difficult task. Questions arise from time to time, for instance, whether an action can be based on a single sentence written by a team member or on a general description lacking in specific detail. More frequently the difficulty is created by conflicts in the evidence regarding a particular matter. This sort of dilemma was considered above. At

present, where such conflicts occur and are not resolvable, the staff tends to recommend provisional approval. Approval or denial of approval is recommended only when there is a clear preponderance of evidence supporting such action.

All of these considerations lead to yet another vexing issue: how many criteria or standards must be violated before denial of instead of provisional, recognition or approval is recommended? Several kinds of considerations must be made at this point—which standards or criteria are violated, the extent and nature of the remediation required, the institution's commitment and resources, and how a standard or criterion is to be applied in a particular instance. Thus, a decision to recommend either denial of approval or provisional approval is likely to depend on a *pattern* or *configuration* of evidence revealing *substantive* violation of one or more standards or criteria.

DUE PROCESS FOR INSTITUTIONS AND STUDENTS: IMPLICATIONS FOR PERIODIC REVIEWS

By Lawrence D. Freeman

Issues concerning due process for institutions, students, and others during periodic reviews arise from the fact that the *Manual* explicitly provides for the possibility of denial of institutional recognition for purposes of teacher education. The act of denying recognition (1) excludes institutions from the group of institutions authorized to prepare teachers, (2) excludes students from the group eligible for certification, and (3) may deprive faculty of their means of livelihood. All these groups—the institution, students, and faculty—may well allege that recognition constitutes a business advantage or a proprietary interest. Thus, the problem becomes one of protecting this advantage or interest while at the same time protecting the state's interest in certificating able and competent school personnel.

At the moment, legal trends suggest that in acting to exclude institutions from accredited or recognized status, institutions may well assert that they have rights or interests to be protected. Currently, there is little case law that provides definitive guidance. But a case involving Marjorie Webster College, a proprietary institution, and Middle States Association, a voluntary accrediting association, may serve to illuminate some of the central issues.

Marjorie Webster College, a proprietary institution, sought to apply for accreditation by Middle States Association, but Middle States refused even to consider the application because it had an eligibility rule excluding proprietary institutions such as Marjorie Webster. The central issue in the ensuing litigation became whether the rule excluding Marjorie Webster was defensible. In this particular case, secondary considerations were (1) whether Middle States' exclusionary rule acted to restrain commerce and (2) whether "education" constituted commerce. The district court found the Middle States rule indefensibly exclusionary; the appellate court reversed this decision finding for Middle States. On appeal, the Supreme Court refused to hear the case. Thus, the decision in this particular case suggests that different criteria are applied in judging the exclusion of institutions from certain benefits than in other sectors of the economy.

The debate, and some of the issues, that this case has raised are illustrated by the commentary of Donald Baker, Director of Policy Planning for Antitrust Division of the Department of Justice:

> I personally think that this case was decided wrongly, even on its own facts. To the extent that it rests on judicial deference to the nature of the underlying subject matter—education—it is a unique little aberration in the law. Secondly, the association's flat refusal to consider Marjorie Webster for membership seems entirely inconsistent with the Supreme Court's decision in *Silver v. New York Stock Exchange*, which required procedural due process for exclusion. Thirdly, the court's analysis of economic effects was inconsistent with the rationale of *Associated Press* and *Silver*. The test has never been whether an enterprise could survive without membership, but whether membership was an "important business advantage." ("Antitrust and the Non-Profit Organization," *Non-Profit Report*, March, 1973, p. 17.)

Some issues remain unsettled. Whether states, for instance, can be found to engage in monopolistic actions in violation of antitrust statutes remains unclear and at this point is mostly a theoretical argument. But if the challenges being raised concerning voluntary accreditation are any guide, challenges of program approval decisions based on broad due process considerations can be anticipated when actions to exclude are taken.

William Kaplin suggests that in situations analogous to program approval:

> Procedural due process would minimally require administrative regularity, reasonable relationship between the decision-making process and the stated purpose of the agency and opportunities for the affected parties to present their views in a manner commensurate with the interest which they have to protect. To ensure due process in accrediting, established standards and procedures should be followed, schools should be given adequate notice and explanation of adverse actions, an opportunity to show cause why it should not be taken, and the opportunity to obtain a hearing and (should adverse action be taken) on appeal; the accreditation body should be free to act without conflict

of interest, pursuant to the evidence and in accord with established standards and procedures. (Quoted from Orlans, *et al.*, *Private Accreditation and Public Eligibility*, 197 , Vol. I, pp. 177-78.)

It is reasonable to assume that students do enroll in institutions and their programs because of state recognition or approval. Counselors at local high schools and two-year colleges base their advice to students upon information provided by the state education agency. Whether that information is intended to be used as such or not, state recognition and approval is interpreted as endorsement of institutions and programs.

Students who enroll at recognized institutions and in approved programs expect to graduate and obtain professional certification by entitlement. If institutional recognition or program approval is withdrawn while a student is enrolled, that student is confronted with the potential loss of a means for securing a livelihood. Because current procedures do not provide due process for those interested students, their rights may be abridged by an *ex post facto* application of authority.

Arguably, faculty may also have rights. Individual faculty members may have elected employment because of state recognition and approval. As in the case of students, faculty members thus have a proprietary interest in their affiliation with given institutions. If recognition or approval is suddenly withdrawn, the immediate earning power of a faculty member is threatened. Further, the effect on the individual's professional reputation could be damaging to the extent that their earning potential is reduced. Unless due process is accorded, the individual rights of these persons may be violated.

Recent student protection cases, Peter Doe and its progeny, as well as the policy research concerning student protection (including prospective teachers) being undertaken by the American Institute of Research for the Fund for Improvement of Post-secondary Education, suggest that in the case of adverse action, students as well might have proprietary rights at stake since they have enrolled and pursued their studies in good faith and with the blessing of the state. The principle that the catalog under which the student enters represents a contract seems to suggest that if the institution truthfully represents itself and its programs as recognized and approved, denial of recognition

or approval may need to be implemented so that present students are allowed to finish their work.

The extent of faculty "rights" is unclear but the United States Supreme Court's ruling in *Board of Regents v. Roth* may mean that individual faculty have a right to a hearing where his "reputation or integrity" is at stake or where a stigma or other disadvantage foreclosing the "freedom to take advantage" of other employment opportunities may result.

The issue then becomes balancing institutional, student, and perhaps faculty interests and rights against the interest of the state in certificating competent and well-prepared school personnel. Due process considerations are obviously *not* the only considerations; compliance with standards and criteria as expressions of state interest are controlling and take precedence.

Caution with respect to due process, however, may be critical, given the character of present technology available for assessing institutions and programs. In 1972, Jerry Miller and Frank Dickey in *A Current Perspective on Accreditation* observed that while accrediting practices vary considerably they share common features:

> Lacking adequate indices and proven techniques of measurement, and occasionally lacking adequate concepts of educational effectiveness or excellence, accrediting agencies have been forced to rely more than is ideally desirable both on personal judgments which are fallible and on quantative factors which do not always have a direct or proven correlation with excellence.

What all of this means is that the program approval process in Illinois must concern itself extensively with both *procedural* and *substantive* due process. Procedural process is provided for in the *Manual* adopted in March, 1975. Substantive due process, essentially, has to do with protection from arbitrary, capricious, and irrational actions. While not precisely defined, substantive due process has to do with fundamental values that emerge when individual and governmental interests compete.

In the context of program approval, the questions associated with substantive due process can be summarized as follows: Is the character of the available evidence sufficient to render the contemplated action (approval, denial of approval) rational, fair, and consistent with stated policy?

RESEARCH LAGS AND INSTITUTIONAL BARRIERS AFFECTING COMPETENCY-BASED TEACHER EDUCATION IMPLEMENTATION

****

REPORT OF THE SPECIAL LEGAL ISSUES CONFERENCE OF THE MULTI-STATE CONSORTIUM ON PERFORMANCE-BASED TEACHER EDUCATION

By Sheila Huff

Sheila Huff is a Research Fellow at the Educational Policy Research Center, Syracuse Research Corporation—a group which has engaged in federal level policy analysis since its founding in 1967. Ms. Huff is currently a case study team leader in the NIE sponsored evaluation of Title I, ESEA Administrative Practices. Her graduate training was in philosophy with an emphasis on the social sciences and education. She served as Research Associate and Acting Director to the Task Force on Competency-Based Education sponsored by the Fund for the Improvement of Postsecondary Education and has authored several papers and articles on education, work and the law.

RESEARCH LAGS AND INSTITUTIONAL BARRIERS AFFECTING COMPETENCY-BASED TEACHER EDUCATION IMPLEMENTATION

By Sheila Huff

Prior to 1973, I studied fair employment practices cases following the Supreme Court *Griggs v. Duke Power Company* decision (401 U.S. 424, 1971) in an attempt to discern the possible implications of these cases for education and the education/work nexus.[1] My interest in and attention to both employment selection and licensing litigation continues but does not approach thoroughness nor the special understanding of the legal scholars I've come to depend upon—David White and Michael Rebell. Both thoroughness and specialized theoretical understanding of the law will be necessary to deal with many of the pressing problems and questions you bring to this conference. If I have anything special to contribute, it would have to do with my rather accidental assignment to several discrete federal policy analysis projects at the Education Policy Research Center in Syracuse (EPRC/Syracuse): one on statewide accountability, one on the implications of the *Griggs* decision for education, and one on the problems and prospects of competency-based education. They are to be sure integrally related, but it was quite by accident that they were merged in my thinking.

For the past two years I have been primarily involved with competency-based education. Thomas Corcoran—then with the EPRC/Syracuse and currently with the Fund for the Improvement of Postsecondary Education—and I assembled a task force which set out to discover what competency-based education was all about and what its problems and prospects might be. Since there were other task forces attending specifically to competency-based *teacher* education, we asked Fred McDonald of Education Testing Service (ETS), Director of the National Commission on CBTE, to join our team which would focus on professional, occupational and liberal education.

---

[1] Reported in Sheila Huff, "Credentialing by Tests or by Degrees: Title VII of the Civil Rights Act and *Griggs v. Duke Power Company*," *Harvard Educational Review*, Vol. 44, No. 2 (May 1974).

Among the many things we discovered was that outside of teacher education there was no one model and no stipulated definition of the meaning of competency-based education. Each of the institutions with up and running CBE programs bore some resemblance to the others, but we could discover no essential commonalities. In CBTE, however, there were attempts to stipulate a model and there were proponents for the movement. Outside of the CBTE movement, competency-based programs in other areas seemed not to take their cues from the developments in teacher preparation. They were engaged in rethinking their institutional missions for a variety of reasons, and in the course of doing so, began clarifying goals and asking whether the instructional content and processes in their institutions were suitable. The language and rhetoric of the CBTE movement seemed appropriate and the special focus funds for competency-based education programs of the Fund for the Improvement of Postsecondary Education undoubtedly influenced their adoption. Several ideas associated with CBTE were recreated and given different labels. Some ideas were directly adopted. The greatest difficulties were with the notion of "behavioral objectives" which most faculty members outside schools of education choke upon. Phrases such as "criteria for identifying learning outcomes" or "indicators of the attainment of educational goals" were more palatable.

If the CBE and CBTE programs had any one thing in common, it was the idea of discovering what competence in various adult roles, including occupational and professional roles, is about and using that information as benchmarks to guide the design of their programs.

Perhaps the first disappointment came when it was discovered by CBE practitioners that there was precious little help to be had from the research and development community on the realities of competent practice in employment settings and still less on what it might mean to be a competent adult in American society. Furthermore, unlike CBTE with its National Commission to organize and coordinate prospective research and development, no such projects were slated for other professions and occupations.

Another disappointment was in the area of assessment of student progress. The expectation was that consultants—testing experts—were there and prepared to venture into unchartered waters—to phase out inappropriate paper and pencil tests, to improve upon the more appropriate ones, and to develop performance measures for new or hitherto neglected educational

goals. Instead it was discovered that with rare exception[2] the testing industry was quite locked into traditional approaches. As a result, most of these institutions began developing new approaches on their own. The result has been some refreshing innovations and the resurrection of some very old techniques. These emerging assessment systems attracted the attention of the testing industry and there is now developing a supply of thick "how to do it" manuals. However, these new methods are highly subjective, for the most part, very costly, and time-consuming. Paul Pottinger, in a paper prepared for the National Institute of Education (NIE),[3] analyzes these problems and suggests how they might be overcome.

The third disappointment was in the area of adult learning theory. When educators were quite clear about what it was a student should learn and how the student would be assessed, they wished to know what kinds of learning experiences would be most efficacious—what content and what process would assure the development of particular kinds of competence. If research has little to say to the previous questions, it has even less to say to the latter.[4]

There were, to be sure, further problems—not the least of which was convulsive change or development under the pressure of outside funding and students on the doorstep. I suggested to Bud Hodgkinson (then consultant to many of the CBE programs we were studying) half in jest, that a physical examination of faculty and administrators in these new programs ought to be given before the conversions and periodically thereafter. We were impressed with the volume of Kleenex tissues in use and always traveled well supplied with Vitamin C tablets. It seemed that changes were occurring far too rapidly

---

[2]Numbering among the exceptions were Paul Pottinger and David McClelland of McBer and Company; Fred McDonald, Warren Willingham and John Valley of ETS; Ruth Nickse of Syracuse Research Corporation.

[3]Paul Pottinger, "Comments and Guidelines for Research in Competency Identification, Definition and Measurement," prepared for the National Institute of Education, Contract NIE-400-75-0036, June 1975. Available through EPRC/SRC, Merrill Lane, Syracuse, New York 13210

[4]See David C. McClelland, "Pedagogy and Competency-Based Education," prepared for the National Institute of Education, Contract NIE-400-75-0036, June 1975. Available from EPRC/SRC, Merrill Lane, Syracuse, New York 13210.

for the general welfare of the idea of competency-based education. I concluded that competency-based education was but a distant ideal—something to be in part deduced from the realities of practice and in part created out of a concern for the future . . . someday when we understood reality better and had a clearer understanding of where we wanted to go.

Observing the competency-based education scene, it could be said that *either* "there are several different CBE models in evidence across the country" *or* "there are no CBE programs to be discovered as yet."

But it is there to be sure—the competency-based movement—forging ahead, being mandated and otherwise encouraged for a wide variety of reasons: (a) on the merits of the ideal model, (b) to attract students in a tight market, (c) to capture educational development funds, (d) as an incentive to stimulate self-examination in higher education, (e) as a means to achieve greater accountability ("make educators tell us what they are attempting to accomplish and when they fail to keep *their* promises, down their dollars").

Why is there a push to convert to competency-based teacher education before conducting the research which may begin to satisfactorily answer such questions as "what teacher behaviors or characteristics effect what student outcomes?" and "how can schools of education promote these behaviors or characteristics in their graduates?"

Among competing possible justifications, one strikes me as having some merit. That is, it might be a good idea to have schools of education geared up to readily avail themselves of promised research findings. It may be worthwhile to proceed now to modularize instruction, to clarify objectives, to set new decision-making structures in place, to develop more creative approaches to assessment *if it is the case that students are being at least as well prepared as they were in the past.* But, of course, we can have no sense of conviction that these changes represent *reform,* since we have no sound evidential basis for such judgments—few or no research findings to temper or mediate judgments of authorities in the field of teacher preparation as to whether CBTE, in particular settings and in general, represents an improvement or a setback.

So, "what is new?" you ask. Perhaps what is new is that never before did people expect so much from the scientific community and have so little

faith in the judgment of experts; and certainly never before did we have a law such as Title VII of the Civil Rights Act of 1964 and knowing, willing litigants to trouble your minds. In the other papers being presented here and in our discussion sessions, we will be exploring the interrelationship between competency-based teacher education and Title VII, so I shall not go into details here.

I do not wish to be interpreted as claiming that the structural, procedural, and instructional changes in teacher preparation occurring under the rubric of CBTE should not precede the state of the art in research that would help define the dimensions of competent teaching practice. My judgment is suspended on that matter at this time. I see advantages and disadvantages. I wish simply to point out that the vocabulary and rhetoric of CBTE raises false expectations, and I worry that the movement might be declared a disaster area before there are any CBTE programs that begin to approach an ideal model. What is being done could be done using language that does not raise false expectations and does not so radically misrepresent reality—thereby reducing the above risk.

Returning to the question "Why put the cart before the horse?", the possible value of promoting changes in teacher education in advance of supportive research findings can be seen in the following example of a case where research outpaced other changes: the Health Services Mobility Study (HSMS) sponsored by the Manpower Administration of the Department of Labor.

The Health Services Mobility Study addresses acute labor shortages in health fields. The purpose of the project was to analyze the tasks performed by medical personnel in radiology, nuclear medicine and ultrasonics. Eleanor Gilpatrick and Christina Gullion conducted the study.[5]

Dr. Gilpatrick, a labor economist, wished to construct career ladders in these job categories that would be based upon similarities and differences in various tasks in particular areas—jobs being delineated or defined and

---

[5]Christina Gullion and Eleanor Gilpatrick, "The Design of Curriculum Guidelines for Educational Ladders using Task Data," Working Paper No. 11 (July 1973), Health Services Mobility Study, Contract No. 82-34-69-34, Manpower Administration, U.S. Department of Labor.

created on the basis of areas of expertise and skill required and their progressive order of difficulty. Information from task analyses was used to construct career ladders and bridges (career ladders applying to movement from one level to the next, and career bridges applying to movement from one occupational field to another). The data was also used to develop curriculum guidelines for educational institutions—to be used in the construction of educational ladders to parallel career ladders on the job. The idea was that it should be possible for people to have access to continuing or recurrent education which would enable them to progress along career ladders or to change to other health occupations with overlapping knowledge and skill requirements. Scarce resources of health educational facilities and personnel would not be wasted in redundant training. Students would not need to sit through courses, the content of which they had already mastered. Job mobility would be increased and worker alienation decreased.

Gilpatrick's project, in all its phases, was influenced and refined by workers themselves, hospital supervisors, professionals on teaching faculties and representatives of professional associations, including the American Medical Association. The results and recommendations of the study were commended by all of these groups as sound and reasonable. All interested parties had an opportunity to refine the task descriptions in light of their judgment of what comprised the best of practice in the observed areas. The research project was well-conceived in a broad socio-economic framework.

At the moment, it is questionable whether the work of the HSMS will be used to fulfill the purposes for which it was funded. Why? Because the primary reason for conducting the research was that it be used to implement reforms in licensing and in education for medical occupations and professions. No one is using the work currently. Perhaps a few professors revising a course here or there may use it to good advantage—but that hardly justifies the cost of the project. To fully exploit it would require thorough, coherent, well-planned systemic changes—that is, efforts in particular settings that would involve hospitals (employers), colleges and universities, and state licensing agencies. The Department of Labor, National Institute of Health, Fund for the Improvement of Postsecondary Education, and the National Institute of Education—among interested funding sources—would have to cooperate and eradicate jurisdictional barriers to make funding possible. I shall outline the work to be done.

The occupational licensing structure is out of phase in all states with the proposed career ladders. This gives the hospitals pause. They hesitate to allow people who can perform work beyond the level at which they are licensed. In fact, to do so is illegal.

From the perspective of the worker, what is most often required to move along a career ladder is a return to a rigid, lock-step formal education program that takes no heed of the abilities acquired on the job and often requires full-time attendance.

From the perspective of educational institutions, the entire incentive structure conspires against utilizing the HSMS materials to create educational programs parallel to career ladders. First, of course, there is no incentive to create programs to reflect non-existent career ladders. Second, there would be great risk to students, even if the career ladders did exist, if the occupational licensing examinations and prerequisites were out of phase with the new programs. Third, there are no declining enrollments in these areas to encourage change.

And could one go to the state licensing agency and propose reforms? Why should they bother with such a tedious and politically delicate task when there is no strong pressure for change?

My conclusion is that unless we work simultaneously on all three areas of reform—the employment setting, occupational/professional licensing, and educational programs—no significant change will result. Moreover, research and reform must operate in tandem.

This should make you feel somewhat better about the competency-based teacher education movement. Research does lag seriously; reform zeal has outpaced it. But you do perceive the broad outlines of what you are about and you do have your commissions and consortia directed at policy analysis, development and planning. Perhaps one day you will even have a lobby in Congress to allocate more funds for basic research on teacher effectiveness and adult learning theory.

Given the story of the Health Services Mobility Study, perhaps now you can see the rational appeal of litigation under Title VII as a prod to reform. Unfortunately, litigation is occurring not in the precious few areas

where research has outpaced changes needed to implement it, but rather in areas where research has not yet uncovered significant dimensions of competent performance–specifically: teaching and school administration. This is unfortunate, because under the gun of litigation at the present time in the area of teaching and school administration, the status quo might be reinforced *or* make-do, hastily conceived alternatives may be adopted. Both possibilities would reduce the demand for and investment in sound research–not to mention undermining confidence in the CBTE movement and possibly undoing what has been accomplished.

In the meantime what is to be done until we have research findings to support educational policies?

1. What are reasonable and affordable interim alternatives to current practices which do have an adverse impact upon populations protected by the Civil Rights Act and practices which are otherwise contrary to our desires for justice and improving the quality of education?
2. What are reasonable interim measures for evaluating competency-based teacher education programs?
3. What are reasonable interim means for evaluating candidates for teaching positions and teacher performance?

These are imposing policy questions which you confront, and, hopefully, we can begin to address them and unearth others over the course of this conference.

REPORT OF THE SPECIAL LEGAL ISSUES CONFERENCE OF THE MULTI-STATE CONSORTIUM ON PERFORMANCE-BASED TEACHER EDUCATION

By Sheila Huff

Introduction

Although it is not obvious from our affiliations, the consultants on the panel, together with Paul Olson, have been working informally over the past year following research directions recommended by the Study Commission on Undergraduate Education and the Education of Teachers directed by Paul Olson. David White, Staff Attorney with the Childhood and Government Project at Berkeley, has made important contributions to these efforts. Unfortunately, he was unable to join us in Denver. Together we are in the process of developing a proposal to establish an institute which would examine occupational licensing, professional certification and related issues. We looked upon this conference as an opportunity to share some of our thinking to date on these issues and to begin clarifying the parameters of our institute's inquiries with the assistance of the conference participants.

Although the planned institute would not limit itself to an exploration of licensing in the teaching profession, it is an excellent place to begin our work. Among occupations, employment practices relating to school personnel seem to be front and center in the arena of the courts, and the competency-based teacher education movement reflects a strong commitment to re-examine standards of professional preparation and certification. Moreover, whatever clarity, depth and breadth of understanding we might achieve with respect to teacher licensing should have significant bearing upon licensing in other areas which face comparable problems.

Attending the conference were staff members of state department of education teacher certification departments, teacher union representatives, faculty of schools of education, members of the research and development community, legal counsel to state departments of education, lawyers representing the interests of parents and students—a rich array of individuals with

a stake in decisions regarding the teaching profession. They came to the conference seeking a better understanding of the implications of Title VII, Civil Rights Act of 1964, for competency-based teacher education (CBTE) and certification. We came wishing to share information and ideas and to determine whether the policy questions we considered important to address were likewise considered important by the participants. That we did not have ready answers to all the policy questions the participants and we, ourselves, raised produced some frustration on both sides. Nevertheless, all agreed the time was well spent and much was learned all around.

This report is an attempt to articulate the questions and to capture the rich dialogue of the conference so that it may be shared with interested parties not in attendance. For the consultant panel and participants, the report represents a point of departure. We now have a clearer understanding of the policy questions surrounding competency-based teacher education and a better understanding of the context in which they will be answered. Developing sound answers to policy questions raised at Denver will, we hope, be on the agenda of future joint meetings of educational decision-makers, lawyers, researchers, policy analysts, teacher representatives and educators.

Conference Moderator: ***Paul Olson***, Director, Study Commission on Undergraduate Education and the Education of Teachers

Conference Coordinator: ***Betty Levitov***, Study Commission on Undergraduate Education and the Education of Teachers

Consultant Panel:

Larry Freeman, Director, Teacher Education Program Approval, Illinois Office of Education, Springfield, Illinois

Sheila Huff, Research Fellow, Educational Policy Research Center, Syracuse Research Corporation, Syracuse, New York

Paul Pottinger, Director, Institute for Competence Assessment, McBer & Co., Boston, Massachusetts

Michael Rebell, Attorney, Rebell, Krieger, Fischbein & Olivieri, Attorneys at Law, New York, New York

I. CONTEXTUAL SKETCH OF CBTE MOVEMENT AND QUESTIONS GENERATED

Most of the discussion of features of the CBTE context arose during the course of the conference as fragmented, elliptical comments. Because of the importance of understanding the context within which policy issues relating to CBTE will be resolved, the author of the report has expanded somewhat upon items from the original discussion in Part I. Time did not permit further elaboration.

The Accountability Movement

The accountability movement may be viewed as the contemporary translation of concerns for social justice, for quality, and for efficient and effective management in the domain of education. The primary intent of accountability legislation is to assess the outcomes of schooling in relation to public resources allocated to education.

In a detailed analysis of the accountability movement as of 1973, Maureen Webster notes that the push for accountability is growing and creating changes in the context for decisions about educational policy. She attributes this development to several trends of the 1960's:

- –The provision of virtually universal public schooling through high school allows and impels an increasing shift in attention from responsibility for access to schooling to responsibility for its outcomes.
- –Massive federal involvement in education has contributed to rising expectations and increasing awareness of disadvantage on the part of those with substandard schooling.
- –The ambivalence of several major publicized evaluations of the impact of massive subsidization has eroded confidence that large infusions of tax dollars will result in substantial improvements in the measured performance of schooling.
- –The costs of schooling are increasing together with the resistance of taxpayers to the rising curve of expenditures on education.[1]

According to the 1974 figures of the Cooperative Accountability Project:

–18 states had enacted accountability legislation pertaining to state testing programs,
–16 states had enacted accountability legislation pertaining to management systems,
–13 states had enacted accountability legislation pertaining to professional school personnel evaluation,
–2 states had enacted accountability legislation pertaining to performance contracting, and
–3 states had enacted accountability legislation relating to performance-based school accreditation.

Most of the states enacted legislation under several of these categories. In total, between 1963 and 1974 thirty states had enacted seventy-three laws containing approaches to accountability.[2]

The first wave of accountability legislation related to statewide testing/assessment of student performance. It deals at varying levels of specificity with what is to be evaluated, at what level, with what target groups, when, and how. Webster points out that whereas in the past testing was used primarily for the guidance of students, contemporary accountability legislation emphasizes a different primary function for testing programs which is the furnishing of information for state-level decisions about the adequacy of schooling. The second wave of accountability legislation involved the evaluation of professional school personnel, with thirteen states having such provisions as of 1974.

Since the idea of relating pupil outcomes to teacher performance has considerable force and tenacity, the fear among professional school personnel is that test scores of students will be *misused* for evaluation purposes and teachers and administrators unfairly blamed for poor performance of students.

Accountability pressures, coupled with the traditional concern of state education department officials for improving the quality of the teaching force, led the conference participants to ask themselves and the consultants:

1. *How can the performance of teachers be more fairly, legally, and adequately evaluated?* or *How should one evaluate teacher competence?* (Policy questions are italicized and numbered consecutively throughout Part I. They are of the form: "What *should* we do?"

CBTE as Institutional Accountability

State Education Department officials are expected to administer the broad policies contained in statutes, regulations, and the directives of legislatures, state boards of education, and chief state school officers. Broad policies must be converted into specific regulations. A reasonable hypothesis is that state policies of promoting, encouraging, or compelling the adoption of a competency-based model by teacher preparation programs have arisen in the course of converting broad accountability directives into specific state education department policies.

Educational accountability, broadly construed, calls for an examination of all aspects of the system—*the goals of schooling, pupil outcomes, teacher performance, the objectives of teacher preparation programs, admissions and evaluation standards of teacher preparation programs, certification standards, the allocation and utilization of resources, the utilization of R & D findings,* etc. Ideally, designing and obtaining approval for a particular CBTE program forces an examination of these and other factors and an explicit reporting of decisions and plans deriving from these considerations. These formal statements, which are the basis of program acceptability "as CBTE programs" in several states, set the stage for accountability of teacher training institutions.

Many questions were generated in focusing on this aspect of the context of CBTE. A second pressing policy question which emerged at the conference was:

2. *How might the states best assure quality teacher preparation programs?*

Related questions are:

3. Does more accountability indeed lead to better program quality?

4. Might these agreements between state education departments and educational institutions (program approval) be treated like contracts, and:

    a. If an institution does not 'honor its commitments,' can the state put it out of business, and would the courts back up the state?

    b. Might a former graduate sue the institution for failure to prepare him/her as agreed?

c. Might a school district sue an institution for recommending the credentialling of an inept student? and

d. Might a student taught by a graduate of a CBTE institution sue the state for allowing that institution to certify an incompetent teacher?

5. Can agreements be treated as contracts in the absence of empirical validation of standards for determining institutional performance?

CBTE as Rational Management

Another interesting aspect of the context of CBTE is the current intensity of the trend toward greater and greater rationalization of work processes --a trend with us since the earliest stages of the industrial revolution.[3] CBTE, the rational equivalent at the institutional or program level of behavioral objectives at the classroom level, reflects this trend. 'Management by objectives,' first used in the business community, is now a household term in certain educational circles. A close look at what is involved in an institution's adoption of the competency-based education mode reveals its debt to 'management by objectives' and other management models such as Planning–Programming–Budgeting System (PPBS), Uniform Accounting System (UAS), and Management Information System (MIS).[4]

Hand in hand with intensive rationalization is the demand CBTE places upon the scientific community. There are calls for new and improved measurement instruments; research on learning of adults as well as children, research on dimensions of competent performance, and so forth. Educators want to be absolutely clear about what they are doing, and to be assured that what they are doing represents the best of practice as scientifically confirmed.

The demand for research and development findings in these domains far outpaces the supply, and there is no assurance whatsoever that the level of funding of this research and development will be increased. Furthermore, there is no assurance that the research and development effort in demand would be coherently organized and effectively coordinated to assure its quality and relevance were adequate funds available.

The other side of this coin is a crisis of confidence in the judgment of those traditionally counted as educational experts. Policy-makers who depend upon experts for advice seem to be saying "We must await scientific confirmation before holding forth any conviction about what we are requiring or what we are doing."

Accompanying crises in confidence are challenges to the authority of professional educators and policy-makers to decide: (a) what the schools should be teaching, (b) in some cases, the manner in which it will be taught, and (c) the criteria that will be used to select and evaluate school personnel. Federal and state agencies have begun, over the past decade, to stipulate the institution of parent advisory councils as a condition of funding and to demand their 'meaningful participation' in educational decision-making at the elementary and secondary level. Also, due perhaps to unionization, classroom teachers are increasingly involved in state-level educational policy decisions.

Although 'involvement' and 'meaningful participation' does not mean control, it does represent entry and influence in the domain of traditional policy-makers and their advisors, and some perceive this as a first step toward community or union control of the schools. Some conference participants hailed this possible development as the best means available to improve the quality of schooling. Others damned it as a development that could destroy the public school system. Whatever the attitude, it was considered a significant development to be taken into account in CBTE planning.

In keeping with this trend toward broadening the cast of educational decision-makers, participation of parents and teachers has a place in the planning of CBTE programs in several states. Questions about the legality and advisability of this participation arose together with other questions dealing with the highly rationalized nature of CBTE, the lack of an adequate scientific foundation, and the crisis of confidence in experts:

6. Does a more highly rationalized process lead to a qualitatively better product? Specifically, do the activities generated by converting to CBTE enhance the probability of developing teacher preparation programs of better quality?

7. Are unrealistic expectations being placed upon the scientific community? Might CBTE be construed as 'false advertising' or 'fraud' in the

absence of a sound scientific basis for requirements? Could we (state education departments) be sued?

8. *Until the scientific community meets the expectations of CBTE* (or comes closer, or gives up), *what rules should govern judgments involved in educational decision-making? What would be the criteria for trustworthy evidence?*

9. How have research findings entered into policy decisions in CBTE? *What should be the place of such findings in CBTE policy decisions?*

10. What is the basis of state authority in educational decision-making? *What should be the basis of state authority to maximize the probability of improving the quality of schooling and the preparation of teachers?*

11. How has authority over educational decisions been delegated in the states? *How might the authority structure be revised or resuscitated to maximize the probability of improving the quality of schooling and teacher preparation?*

12. Is the crisis in confidence in the judgment of traditional experts justified? *What should be the role of traditional experts in CBTE decision-making? How much weight should be given to their judgments?*

13. *What educational decisions should parents, community representatives, teachers, and students influence? Which should they control?* What has been their role in CBTE decision-making? *What should be their role in CBTE decision-making?*

Cultural Pluralism

Another important aspect of the context of CBTE are dilemmas of cultural pluralism. Insofar as the goals of public schooling are anchors of CBTE, and insofar as goals are very much up in the air and the subject of controversy, CBTE is not secure in its moorings. Definitions of competence derive from conceptions of practice. We are in a period of rather intensive re-examination of the aims of education. The old metaphor of the melting pot as an ideal is being reconsidered, and we wonder how cultural diversity might

coexist with certain commonalities in education that enable people to cooperate in building a better society and a better world. Recognition of these societal tensions generated some difficult questions:

14. *How do we cope with ever-changing sets of educational goals? What do we do when various groups cannot agree on a set of goals? How much homage is to be paid to articulated goals of schooling in preparing teachers?*

15. *How do we fit various prescriptions for dealing with cultural diversity into CBTE? Who has authority with respect to goals? How should we make the tradeoffs between general and culturally specific training and licensing?*

16. *What would be the characteristics of an ideal licensing system that would fit a schooling system which provided a common education as a core and a broad range of options beyond the core to accommodate cultural diversity and special interests? What should be handled through local hiring and promotion practices, and what should be relegated to the domain of teacher licensing? What would the state's role be with regard to local hiring practices? How should 'relicensing' or in-service evaluation interrelate with local prerogatives?*

CBTE and Fair Employment Practices

Last but not least—and probably foremost—in our examination of the context of CBTE is Title VII of the Civil Rights Act.

Somehow, somewhere, thoughtful people concluded that inequalities of employment opportunities in the United States were not *necessarily* the result of *intent* to discriminate against particular populations. Rather, built into the very fabric of institutions or practices affecting employment selection were rules, processes, structures and standards which, as they *legitimately* or *inevitably* discriminated, did so disproportionately with respect to particular populations such as minorities and women.[5] Insofar as these could be justified—in particular cases—in terms of the requirements of adequate job performance, they were deemed rational, fair, justifiable, 'a matter of business necessity,' and therefore, allowable. Insofar as they could not be so

justified, they were deemed instances of 'institutional racism' or 'institutional sexism.' Those responsible for imposing these rules, processes, structures and standards were to be held accountable for their (unjustifiable) adverse impact upon minority populations and women, despite the absence of any intent to injure these populations. Thoughtlessness or ignorance would not remove the burden of responsibility.

Of course, once those imposing these rules, processes, structures and standards understand that they operate to discriminate unjustifiably, continuation of these institutions or practices would no longer count as 'unintentional injury.'

One can readily see how difficult it would be to sort out the innocent from the guilty and how this difficulty might lead one to refuse to acknowledge differences–which is the course that has been taken by the U.S. Equal Employment Opportunity Commission (EEOC) in its *Guidelines on Employee Procedures.*[6]

The making of such a distinction (between intent or no intent to discriminate) on the part of the courts has not been the usual experience of employers who have been involved in Title VII litigation. *Absent* a finding either way, employers who could not demonstrate the job-relatedness of certain employment practices which resulted in 'adverse impact' on populations protected by Title VII have been required by the courts to pay damages and otherwise make compensatory adjustments.[7]

An interesting post-conference development, in a case where discriminatory intent had not been established, is that an employer was held liable for damages to a man whose promotion, after an out-of-court Title VII settlement, was 'adjusted.' Daniel McAleer, an AT&T employee, was bypassed in a promotion decision. To make amends for past adverse effects of promotion procedures and policies (AT&T never admitted to an *intent* to discriminate), AT&T promoted, instead, a woman with a lower performance rating and less seniority.[8] The logic of the courts decision in this case was that "since McAleer had no responsibility for AT&T's past sex discrimination, it is AT&T rather than McAleer who should bear the principle burden of rectifying the company's previous failure to comply with the Civil Rights Act of 1964. An affirmative award of some damages on a 'rough justice' basis is therefore required and will constitute an added cost which the stockholders

of AT&T must bear."[9]

At the time of the conference in Denver, the lower courts, in deferring to the EEOC *Guidelines*, were acknowledging the precedent established by the Supreme Court in *Griggs v. Duke Power Company* (401 U.S. 424, 1971) not only in cases brought under Title VII, but in cases brought under the Constitution's due process guarantee and other provisions such as the Civil Rights Act of 1866. On June 7, 1976, the Supreme Court, in a 7 to 2 decision,[10] ruled that in cases not brought under Title VII of the Civil Rights Act, to prove a violation of the Constitution's ban against racial discrimination, the Court must prove 'racially discriminatory purpose' in addition to a 'substantially disproportionate' burden on one race of particular employment practices. "The court found essentially that the rules for Title 7 cases did not apply to these other provisions."[11]

From these and other developments it may be concluded that the legal context of CBTE is, at the moment, somewhat unpredictable.

It was primarily a concern with the implications of Title VII for state education department activities relating to CBTE that brought us together at Denver. There have already been suits involving educational employment decisions in several states, and there is every indication of more to come. Participants wondered if they could be sued, for what precisely, how they might defend themselves, and how they might prevail against the likelihood of such suits. Their questions included the following:

17. Would or could Title VII be applied to teacher licensing? Isn't it applicable only to hiring and other employment practices—the province of school districts rather than state education departments?

18. We cannot yet establish predictive validity for the measures used in teacher certification, and predictive validity is preferred by the EEOC *Guidelines*. Would the courts be content with a demonstration of content validity?

19. How might we go about content validation that would satisfy a court?

20. Absent a demonstration even of content validity, would the courts settle for a democratic process involving parents, teachers and students

in arriving at standards for teacher certification?

21. Among the professions, is teacher licensing in a less tenuous position with respect to the EEOC guidelines than other areas of professional licensing? (the answer was 'no.')

22. What sort of snares does the program approval approach to teacher licensing hold with respect to Title VII?

23. Given the absence of a scientific foundation for judging the adequacy of CBTE programs, what might be the legal ramifications of a state's not approving or withdrawing approval from a CBTE program? *What kind of evidence could or should one use for a program approval process which would permit one to say 'no' to an institution on some legitimate basis.*

24. Where, legally, is a teacher training institution if it should refuse to certify an individual?

25. *What should be the foundations of an appeal process where students might challenge an institution's or a state's decision not to certify them?*

26. What rights do teacher preparation programs have with respect to state departments of education? Can they reject the CBTE model or aspects of it? How would the courts deal with such a situation?

27. Where would a state stand, legally, if it ordered the dismissal of a teacher in good standing considered competent due to some error in the issuance of a provisional or permanent license? For example, if it were discovered that the teacher had not taken all of the courses required in the state for certification?

28. *What would an ideal certification system look like:*

    a. given current limitations in scientific knowledge,
    b. given an ideal level of scientific knowledge?

29. Subject matter mastery has face validity as a prerequisite to licensing.

Would the courts accept subject matter mastery tests based on careful job analyses?

30. Supposing the goals of education for elementary and secondary schools of a region/locale/state were taken into account in establishing licensing requirements, how could we keep apace of changes? *How should goals of education be taken into account in program planning and program approval?*

31. *What should be the role of the research community in educational policy decisions? How should we treat research findings? What do we do when researchers disagree or tell us the 'n' was not large enough to reach a definite conclusion?* "How large an 'n' do you people need anyway?"

32. What has Title VII done to the definition of "policy"? It used to be the case that if a person in authority in the educational system said something was necessary, it was policy, and it was not questioned by the courts.

33. *What do we do when what is do-able, feasible, is in conflict with the way we believe things should be done? What should we do now to assure that we are not foreclosing on the options of 'best' licensing systems?*

34. State authority in most cases is at the initial entry level into teaching. With declining enrollments and very little turnover in the teaching force, is not the real issue how to make the present work force more competent? *Should we approach the problem of school personnel evaluation through technical assistance to local districts, a second-tier of certification, relicensing? Should pressures be put on schools of education to focus upon in-service training?*

35. What are the primary sources of potential challenge to credentials? *What should our priorities be in making our systems as rational and fair as possible? Should we set priorities in accordance with the order of vulnerability?*

36. When a state legislature passes a law that bilingual, environmental

protection and conservation programs will be instituted in the schools, what are the obligations of state education department officials in charge of teacher certification? The teachers we are licensing will not necessarily teach in our state. There will be teachers coming to our state from elsewhere. *How should this be handled?*

37. Some states have renewal-of-licenses cycles and standards that are not being enforced. *Should they be enforced? How?* What are the obstacles? How might they be overcome?

38. *What should we do about the 'good moral character' prerequisites for licensing?* (The answer was "we need another conference to address that.")

The number and complexity of these questions—especially the policy questions—should be enough to convince the reader that they could be dealt with in only the most cursory manner in the few hours at our disposal. All of them are serious and deserving of the most careful attention. What follows is an attempt to capture the essence of discussions relevant to most of the questions. Some of it is information which stands on its own. Some of it represents tentative suggestions, ideas, opinions, attitudes, perspectives—all meant to be helpful, but far from the final word on the issues raised.

II. DISCUSSION

'Rational' v. Rational: 'Empirical' v. Empirical: and the Courts v. The Research Community

At some point during the conference, after a period of talking past one another, Thomas Green (Syracuse University) pointed out that the researchers, lawyers and state officials present were bogged down in contending notions of 'rationality.' From a contextual analysis of the EEOC's *Guidelines*, one may conclude that its authors construe 'rationality,' as it applies to employment decisions, as 'proceeding according to the rules of scientific inquiry.' The EEOC specifies the particular rules of scientific inquiry to be followed in developing and validating assessment devices. The procedures are rigorous and systematic.

On the other hand, our ordinary language understanding of 'rationality,' and that generally shared by policy-makers, is reasonableness or good judgment in light of available evidence. In most situations, we do not insist upon systematically vigorous standards of confirmation in deciding whether an action is or is not rational.

As noted in Rebell's paper, the treatment of the concept of 'rationality' by the courts is complex and relevant to the issues under discussion at the conference. The courts have traditionally distinguished between those types of claims which would receive 'strict scrutiny' (generally, cases involving certain categories of victims such as racial minority groups, or certain 'fundamental interests' such as voting rights) and all other cases which the court would review in a perfunctory fashion.

In the case of 'strict scrutiny,' the state would have to show that the challenged law or practice was absolutely necessary, that there was no other way to accomplish the result. In cases not receiving 'strict scrutiny,' the courts would invalidate a state statute or action only if there was no possible rational relationship shown between the statute or action and any conceivable legitimate goal or purpose. Rebell pointed out that in recent years this dichotomy between 'strict scrutiny' and rational relationship analysis has begun to break down, and a middle ground of analysis is emerging. In several cases, in applying the 'rational relationship test,' the courts undertook a substantive rather than a perfunctory analysis and held that the requirements

being challenged were not substantially related to the purpose for which they existed. The courts, in other words, 'revitalized' the traditional rational relationship approach and gave it some bite.

It was concluded that it is important to take into account these differing treatments of 'rationality' in order to avoid ambiguity.

Some additional confusion in the discussion stemmed from difficulty with the concept 'empirical.' The scientific community restricts its usage to propositions confirmed in accordance with rigorous scientific procedures–a special class of experience; whereas the ordinary language understanding of 'empirical' is 'based on any kind of experience.' In other words, standards of confirmation differ.

It was noted that the EEOC *Guidelines*, then, may be viewed as an attempt to change the concept of 'rationality' in the courts with respect to Title VII issues (through changing the standards of scrutiny and confirmation)–to base the concept more upon empirical evidence derived from *systematic, rigorous inquiry* and less upon empirical evidence from ordinary experience. That is, with regard to employment selection devices, where their use results in adverse impact upon populations protected by Title VII of the Civil Rights Act of 1964, there is a requirement for a clear showing of 'job-relatedness.'

As pointed out in Michael Rebell's paper, "Validation under the EEOC *Guidelines* is described under two main headings: 'predictive validation' and 'content validation.' Predictive validation is the higher standard, which is normally required of all employers. . . ." After briefly outlining what is involved in predictive validation, he points out that "Although predictive validation is the preferred standard and is required generally in cases involving private companies, in many of the public sector cases which have arisen, the courts have invoked a clause in the EEOC *Guidelines* which permits the use of 'content validation' where predictive validation is presently 'not feasible.' "

In his paper, Rebell also explains developments in the courts which lead him to expect that in cases involving challenges to teacher licensing practices –even if not brought under Title VII–the courts will invoke the EEOC *Guidelines* in looking for a rational relationship between the requirements

of performance as a teacher and the prerequisites of licensure.[12]

Given the following circumstances, his prognosis is that the courts would be satisfied with a showing that tests or other criteria in use have been *content* validated by carefully conducted job analyses:

(a) It is possible to show adverse discriminatory impact of selection devices in almost any particular case brought before the courts (therefore, states can expect the burden of proof to be on their side in challenges brought under Title VII); and, CBTE notwithstanding.[13]

(b) It is not possible at present to demonstrate *predictive validity* for any of the selection devices in use.

(c) No alternative selection devices with greater *predictive validity* and less discriminatory impact are available to substitute.

(d) There is currently no body of scientific data upon which to construct tests with greater predictive validity.

Rebell recommended close attention to two New York City cases, *Chance v. Board of Examiners*[14] and *Rubinos v. Board of Examiners*[15]–involving an additional level of certification–for insights into how the courts might respond in challenges to state school personnel licensure requirements or procedures. He suggested that for licensure purposes, under the circumstances, subject matter tests based on job analyses would have face validity and would probably impress the court. Fred McDonald interjected that studies conducted by ETS revealed that for all its face validity, subject matter mastery explained only 7 per cent of the variance in teacher performance as measured by pupil outcomes.

This point brought to light one rationale behind resistance to *minimal* licensure based on subject-matter mastery tests. That is, in addition to *knowing about* what one teaches, people want assurance that teachers *know how to* teach others what they know. Apparently, it is thought that *some* proficiency with *some* methods–even though one cannot demonstrate which are most effective–is preferable to no training or no proficiency with any methods. The 'approved program' approach to licensure assures, at minimum,

exposure to some methodologies of teaching, *if not proficiency*. Sharpening assessment procedures in schools of education is one way in which CBTE hopes to improve the likelihood of greater proficiency. The frustration, of course, lies in the doubt that the proficiencies promoted are significantly related to teacher effectiveness.[16] This theme of doubt pervaded the discussions.

Apart from the insufficiency of subject matter mastery for licensure or hiring, it is, to be sure, a necessary component of teacher competence, and subject matter mastery tests *are* used extensively as screening devices. Rebell's point was that the courts would, in all likelihood, insist upon a showing of rational relationship between the subject matter of tests or program requirements and the area of teaching for which teachers are being licensed. One basic step in demonstrating rational relationship is to conduct thorough job analyses–the key to acceptable *content validation*. In his paper, drawing upon his experience in *Chance* and *Rubinos* cases, he sets forth key components of a valid job analysis. Summarizing from his paper, they are:

(a) The specific job areas should be clearly defined yet not atomistic. (Paul Pottinger's paper gives insights into how one might avoid undue fragmentation.)

(b) A representative cross-section of persons who have an interest in the selection or evaluation process must have input into the job analysis. (The job cannot be defined exclusively from the perspective of present job holders because of upgrading and updating responsibilities.)

(c) In the case of regional or local licensing in addition to state licensing (as in New York City and Buffalo, New York), local needs should be taken into account by including the perceptions of local policy-makers as well as the 'traditional' employer (e.g., principals, parent advisory councils, etc., as well as superintendents and state officials).

(d) The use of independent outside consultants to undertake the job analysis to prevent monopolization of the process by people with vested interests in particular outcomes.

With regard to possible court challenges, Paul Pottinger pointed out

that most of the examinations that were in use were *not* based on job analyses. They were *not content validated.* The participants asked Pottinger and Rebell if they would outline how to go about thorough job analyses that would pass muster in the courts. Since procedures for conducting job analyses are technical, there simply was not time at the conference to provide a sketch of what is involved. As well as techniques the participants wanted: (a) estimates of the costs that might be involved, (b) ideas on how they might cope with changing needs and goals, and (c) information that would help them in deciding what should be the province of state licensing at the entry level to the profession and what should be the province of local hiring or possibly a second level of state certification. In other words, it seemed to the participants that conducting job analyses might be a good idea.

Paul Pottinger registered some reservations in this regard, however. He questioned the wisdom of investing in job analyses without empirical evidence supporting their validity. To simply determine that tests are related to job analyses without evaluation of the validity of the job analyses, themselves, provides a tautological system which may satisfy the courts' demand for due process but not the scientific demand for better knowledge of what works and why. Since content valid tests are only as good as the job descriptions, attention to the validity of job function analyses must be given higher priority if licensing tests are to be relevant to competence and not merely relevant to rational or due process criteria.

In his paper he pointed out that only construct validation—constructs being tied to competent performance as measured by carefully controlled pupil outcome studies[17]—would sort out effective from ineffective practices; and "is that not, after all, the work at hand if the goal of CBTE is to improve the quality of schooling?" Money and time invested in job analyses would drain funds and attention away from critical basic research on teacher effectiveness—an already grossly underfunded domain of inquiry. Further, he urged against employing the energies of the CBTE movement in playing a 'catch-up game' with the courts. He recommended a 'proactive' rather than 'reactive' stance with best energies being devoted to developing methods of arriving at sound answers to the questions CBTE raises.

A state education department official suggested a compromise: support basic research in teacher effectiveness through enlisting congressional support for NIE's projects in this area. In the interim—which may be decades—

proceed with job analyses beginning where they seem to be most needed. Pottinger noted that adequate technology for conducting job analyses already exists which goes a long way toward meeting both the demands of the courts for due process and the demands of methodologists for constructing valid job descriptions. However, no consensus was achieved on next steps to be taken.

"What," it was asked, "might be the courts' reaction if requirements could not be backed up by systematic job analyses let alone carefully controlled research studies on effective and ineffective practices?" Although the defendants would be in a very tenuous position in light of developments reported in his paper, Michael Rebell's response appealed, again, to the courts' construal of 'rationality':

(a) Was the matter of requirements given the careful attention of individuals the courts would acknowledge as experts?

(b) Can the state education department produce rational arguments in support of the requirements at issue?

(c) Were the decision procedures for deciding fair, with the interests of groups affected by the decision taken adequately into account?

(d) Were there established, operable procedures and guidelines for appeals?

These, he suggested, were the kinds of things to which a judge would pay careful attention in arriving at a decision.

Participants wanted to know how they could prevent suits. Needless to say, the answer was "There is no way to prevent being sued." However, it was suggested that time spent doing the following might very well reduce the incidence of suits and the risk of defeats:

(a) analyzing requirements—revising them where insupportable and substantiating them where reasonable in light of currently available evidence ('best evidence' being from job analyses),

(b) examining and improving decision-making processes, and

(c) revitalizing or establishing sound appeals procedures which eliminate occasions where injuries to parties may be attributed to irrational, technical requirements.

As an example of what he considers a paradigm of an irrational technical requirement, Rebell offered the example of a current client of his who had been teaching chemistry for several years. He was dismissed from his position when it was discovered that he had not had the period of supervised practical teaching required under one (of two) alternatives for New York City licensing (under the authority of the New York City Board of Examiners). In order to qualify for the licensing examination, under Alternative A one had to have completed the required courses *and* a period of student teaching. Alternative B, to accommodate teacher shortages, waived the student teaching requirement

In completing his examination application, Rebell's client inadvertantly checked Alternative A as the category of licensing. He took the examinations, passed, appeared on the Alternative A eligibility list, was assigned to a school, served for three years, had high performance ratings, and was recommended for tenure. A year or two later, in checking through the records, it was discovered that he was not eligible *under Alternative A* to take the examination and that his name should not have appeared on the eligible list *under Alternative A*. He was informed that unless he could prove that he was eligible at the time, under *Alternative A*, his license would be revoked. Of course, he could not because he had not fulfilled the Alternative A student teaching requirement. He should have checked *Alternative B*. His license was revoked and he was dismissed from his job.

Having exhausted other possible avenues for reversal of this decision and reinstatement in his job (that is, the appeals process through the New York City Board of Examiners and the New York State Commissioner's office), he appears to have no alternative but to sue in the federal courts–an alternative he simply cannot afford.

We do not know what the outcome in this case will be–whether it will reach the courts, be satisfactorily resolved outside of the courts, or become another statistic in the accident record of irrational licensing and hiring practices.

CBTE Planning

Defining Competence

The discussion then moved to CBTE planning. When a state decides to adopt the CBTE model, it must decide what, substantively, *will be meant* by 'competence' for purposes of program approval and the certification of graduates. The call is to create an 'ideal type' concept of 'the competent teacher' within specialized areas and in particular contexts. Unfortunately, as pointed out, the 'ideal type' construct cannot at present emerge from well-founded theories of effective practice. New York State had proclaimed that by 1990 it would.[18] At the conference, a representative of the New York State Education Department said, "That was an ideal, a goal, not a promise or a prediction."

Given this lag in supportive research and other considerations brought out at the conference (for example, the pressures to deal more democratically with issues of cultural pluralism and local needs), several states have opened up the planning process to a broad range of participants which have included community leaders, teachers, parents, and students in addition to traditionally recognized experts. Phase I of the process has typically culminated in the production of lists of competencies–some numbering in the thousands–which teachers should manifest.

Many of the conference participants–especially those representing the research community–took a dim view of these procedures *as a means of establishing what teacher preparation programs should attempt to accomplish and state education departments should certify.* Better to rely on experts. (For an argument against this position, see the section on The Courts' Perspective on CBTE Planning on the next page.) However, they agreed to the need for participation of these groups in establishing the goals of primary and secondary schooling.

Delegation and State Interest

Appropos of establishing the goals of schooling, someone asked what the 'state interest' was in such decisions. State constitutions, statutes, and commissioners' regulations frequently specify categories of curriculum for the schools. Furthermore, it was mentioned that if New Jersey is the begin-

ning of a trend, state laws will become increasingly specific about what comprises a 'thorough and efficient education' leaving even less room for maneuverability than at present.

Participants wondered how 'state interest' interrelates with CBTE programs and processes—not simply on substantive grounds, but on procedural grounds, such as *delegation of authority*, as well. Michael Rebell pointed out that this had to be addressed on a state-by-state basis, since state constitutions and statutes differ.

In most states, the constitution has not given the legislature all of the power in educational policy-making. Some powers have been given to constitutionally established boards of education and institutions of higher education. Some of the power of institutions or boards includes the preparation of teachers. For all states, it is apparent that constitutionally appointed bodies cannot do everything. Therefore, they have been given power to delegate policy-making and administrative authority. However, the constitutionally appointed body does retain ultimate control.

In the case of CBTE planning, policy-making authority is often delegated to consortia of community groups, higher education institutions, teachers, etc. with the hopes that the process will result in better standards for competency. Insofar as policy is decided at this level, it must be examined in light of delegation theory. Delegation to this lower level of authority and the widespread adoption of the 'approved programs' approach to certification mean that, *in a classical legal sense*, litigation becomes much more complex and problematic.

In summary, depending upon authority structures and circumstances, delegation will or will not be legitimate. Consideration of delegation, when it involves Title VII and constitutional guarantees, will be very complicated. Whoever retains delegated responsibility and authority will be challenged when standards are at issue. Each state will differ.

The Courts' Perspective on CBTE Planning

Michael Rebell was asked what view the courts might take on the aforementioned participatory CBTE planning process. He outlined some pragmatic considerations which might favorably dispose the courts to the

adoption of these procedures. To begin, he noted that rationality of the procedures and legitimate delegation of authority would not obviate the probable requirement of the courts that the deliberations of these consortia be firmly grounded upon information gleaned from carefully conducted job analyses (in the absence of predictive validations). He thought it likely that the courts would take into account the fact that there are few, if any, better alternatives open to the states at present. The major argument against the most popular one—*continued exclusive reliance on traditional experts*—might, he suggested, go as follows:

> Where minority populations are concerned, considerable evidence can be amassed that points to the failure of traditional experts to effectively deal with minority education. Much of the failure is attributed to a lack of awareness of cultural differences that generate special educational needs. Bilingual education programs and the community schools movement are responses to this failure. Opening up CBTE planning to broader participation that includes minority representation and representation of teachers experienced in dealing with these populations may be viewed as another 'sensible,' 'timely' response to this failure.
>
> Further, insofar as *all* populations are concerned, a vital step in planning is *setting goals*. Since it is not self-evident that in a democratic state traditional education experts should monopolize *this* aspect of planning, including nonexperts in this capacity would, more likely than not, seem reasonable to the courts.

On balance, Rebell's opinion was that: (a) sound arguments can be produced to support the participatory processes of CBTE planning, (b) these arguments would probably impress the courts, and (c) the major concern of state education departments should be to improve the empirical base upon which CBTE consortia develop policies—not simply to protect themselves in lawsuits, but in the interests of sound planning.

Teacher Preparation Programs v. State Program Approval Offices

As noted, it is sometimes the case that teacher preparation programs are

handed long inventories of unweighted competencies drawn up by statewide consortia. These lists are supposed to serve as the foundation of program planning–policy upon which to formulate yet more specific policy. Even where the matter of defining competency is left to individual institutions to determine, they are sometimes required to arrive at their definitions through local participatory consortia. Alone or en masse, in planning CBTE programs, educational institutions are supposed to take a staggering number of factors into account (including their monetary resources and faculty talents), arrive at plans in short order, and live up to them. Several states have instituted CBTE goals, timetables, applications review schedules, and renewal cycles.

The pace and magnitude of change, coupled with self-consciousness about the foundations of the CBTE movement, created acute concerns evident in the questions of participants. It was asked if quality could possibly improve under these circumstances, to which someone responded that things could not be worse than in the past when there was almost no accountability. Unfortunately, as one participant pointed out, the innovation was not being piloted as a social experiment and carefully evaluated as such. Therefore, at present only impressionistic statements can be made about whether it is indeed better or worse. One state education department official said that they realized that in the earliest phase of CBTE in their state, it could be that they would only achieve better accountability, but that over time, they expected that hard thinking and written commitments about what was to be accomplished would yield qualitatively better programs.

As might be expected, one of the participants wished to know if a teacher preparation program could successfully sue the state if, in the judgment of its faculty and administration, the timeliness or requirements were inhibiting the delivery of quality education. Even before this question was asked, state department of education officials worried aloud about the legal grounds upon which they stood in disapproving CBTE or, for that matter, traditional programs.

In the discussion following these questions, Michael Rebell reiterated some of the laws that might come into play given this set of issues. Regarding an institution's suing the state, he reviewed the recent opinion of the Texas Attorney General (Opinion H-197, January 4, 1974), which involved illegal *delegation* of legislative authority to administrative officials. Rebell said that although that opinion was specifically based on an analysis of

certain Texas statutes (the Attorney General ruled they did not permit the State Commissioner to mandate that all teacher education programs in the state be competency-based), the framing of the issues in that situation leads to the broader question of whether a fundamental policy shift such as CBTE may be mandated by an administrative agency in the absence of detailed legislative authorization.

He also mentioned possible legal challenges arising under the First Amendment to the Federal Constitution. "Conceivably, a school that, for whatever reason, does not want to accept a mandatory CBTE directive from a state education department might claim that mandatory CBTE inhibits the ability of the school and its individual faculty members to teach in the manner that they believe best suits their needs, their values and their purposes."

In the case of states refusing to approve programs, the participants wished to know what evidence, in the absence of scientifically derived evidence, would hold up in a court. Larry Freeman suggested that the processes of program approval deliberation involve publicly appointed representatives of 'the public interest in the education of teachers'—that doing so would give more 'clout' to the decisions of program approval offices. Michael Rebell stressed the importance of due process in program *disapproval* deliberations—weighty decisions affecting the power of an institution to grant degrees or affecting the eligibility for certification of its graduates. He suggested that the adversary model of the courts might be most appropriate—testimony of faculty, students, supervising teachers, etc., as well as the usual testimony of 'inspectors' from accrediting agencies and state education departments, being critical evidence to be evaluated in reaching such decisions. Records from such proceedings would be part of the state's evidence in the event of litigation and would rest the state's case on a firm procedural foundation the courts would respect. Moreover, such care in proceedings would probably reduce the likelihood of suits.

A participant pointed out that it was usually the factors of finances and personnel rather than carelessness or obstinacy which led to less than adequate procedures for such decisions. Litigation would be more costly, true, but improving processes would tax already overburdened budgets and staff. As Thomas Green put it, in the press of policy-making and administration, what is rational, in one sense of 'rational,' may not be the best thing to do. There are always values to be traded off—'justice as due process' being

among them. The participants and consultants took this, among other frustrations, home with them.

Licensing v. Employment Selection

One of the participants commented that in his state, given cutbacks in enrollments and in school staff, instructors were teaching subject matter other than that over which they had been examined for certification and hiring, and that as cutbacks increased, this could become the rule rather than the exception. He wished to know what subject matter one tested for under such circumstances. Further, he suggested that the emphasis in CBTE on preservice training and certification was perhaps misplaced–that given a shrinking elementary and secondary education system, perhaps the emphasis of CBTE should be on in-service evaluation and training. This comment led to a whole series of complex questions.

It was pointed out that given a shift in emphasis, many of the most important questions would remain the same. There would still be a problem with criteria for evaluation of teachers and the development of sound pedagogies to enhance the likelihood of improving the performance of practicing teachers; and the legal problems would still involve Title VII as it applies to promotion, tenure, seniority, dismissal, etc. The soundness of evaluation procedures and criteria has been scrutinized in several Title VII and related cases.

Considering what should be the province of licensing and what should be the province of employment practices generated three dilemmas:

(1) the 'mini-max dilemma'
(2) the 'general v. specific dilemma'
(3) the 'common culture v. subculture dilemma'

The 'Mini-Max Dilemma'

"Are we certifying competent teachers, or not?" was one way the mini-max dilemma was stated. In recognition that theoretical knowledge of teaching is insufficient to skilled performance as a teacher, schools of education have incorporated into the curriculum short practice teaching periods where

novices are supposedly closely supervised by master teachers. The principle is sound, but there is widespread recognition of frequent failures of practice teaching to provide quality learning experiences for novices. Therefore, there is a general hesitancy to affirm that schools of education are certifying *competence.* This queasiness was reflected in questions of the participants. "Can a district sue us for sending them a certified 'competent' teacher who proves to be quite incompetent?" "Can parents sue us for turning loose an incompetent teacher on their children?"

Because many educators do not believe that the short term of trial in student teaching is an adequate basis for certifying *competence,* some prefer it to be said that what schools of education are doing is "*maximizing the likelihood* that those who are certified to teach will prove competent to teach" or "*minimizing the likelihood* of incompetence." That, of course, throws into question *competency*-based teacher certification at the *entry* level. Should it be a 'minimal' notion with tougher competency standards left to a later phase of certification? Michael Rebell advocated this approach—'minimum' being subject matter mastery at the entry level. He added that this, of course, would be unacceptable for reasons discussed earlier if not coupled with a sound system of on-the-job evaluation. This would facilitate supplementing the 'approved programs' approach with examinations open to individuals who had not graduated from approved programs—a course of action already taken in some states.

A representative of a teachers' union objected to a two-tiered system of certification on the grounds of unsound criteria for constructing on-the-job evaluation systems and the improbability of locating the requisite number of competent evaluators. Michael Rebell pointed out that *relicensing* involves the same sorts of considerations as a possible two-tiered licensure system and, despite its cogency, is quite as unpopular among practicing teachers and quite as problematic in light of the current weakness of the empirical base for evaluation.

Someone wished to know if minimal qualification should be the only province of state licensure with maximal qualification (*competency* certification) left to employer hiring and promotion practices. Was it not in the interests of employers to hire the best qualified? Needless to say, some participants thought it might not be, given tight budgets and political problems. In essence, their grumbling could be interpreted as expressing the belief that

the public interest in quality schooling requires that state education departments retain the authority to certify *competence.*

The 'General v. Specific Dilemma'

Actually, several dilemmas are encompassed under this heading. One participant described a bind state education departments face. If they are not specific enough in their policies, they cannot be enforced or enforcement becomes arbitrary. If they are too specific, substantive due process or First Amendment 'freedom of expression' and 'pursuit of livelihood' issues arise. "We feel like we are walking a tightrope"–a classic inescapable dilemma of policymakers.

Another critical area of policy decision-making involves general and specific abilities. Teachers are licensed to hold any number of *jobs* within their area of specialization. No one can calculate what specific demands will be placed upon a teacher in a given job. Participants wished to know the implications of this observation for the 'what,' 'when,' 'where' and 'by whom' of licensure. They also wondered how this affects teacher program planning and approval. Should program approval only involve general abilities that would be required given a broad range of jobs in a field of specialization? Are there such things as general abilities? How context bound are abilities?

These questions, of course, *have been* resolved in current systems of licensing. That they were raised reflects the power of the CBTE movement to force a serious re-examination of these fundamental issues. That no one at the conference had ready answers was a source of disappointment to some; but the willingness to confront the tough questions is, indeed, praiseworthy.

The 'Common Culture v. Subculture Dilemma'

Another difficult set of issues, tossed for the most part into the lap of educators to resolve, has to do with cultural pluralism. From the perspective of teacher training, licensure, and employment practices, how shall it be accommodated?

Someone suggested that answers might have to wait upon decisions as to how it is to be accommodated in the schools–decisions reflected in statements of educational goals. Someone else pointed out that it *is* presently

'accommodated' in teacher training, licensure and employment practices, *is* a subject of rather intensive criticism and re-examination, and that there is no reason for the CBTE movement to wait upon decisions from other sources. In other words, the participants seemed to believe there was no alternative to "walking out on a limb" and rejecting the *status quo*.

Rebell pointed out that issues of cultural pluralism are a significant locus of potential challenges to current licensing and employment arrangements. He mentioned problems in making the professional/paraprofessional distinction. The following brief scenario (constructed by Sheila Huff) may help to clarify the point:

> To carry out its bilingual education program involving Puerto Rican children who could not speak English, given a shortage of certified professionals, men and women in the community who spoke both English and Spanish in the Puerto Rican dialect were recruited. They were screened for linguistic ability, experience with children, and interest in teaching. After a short period where their performance was evaluated, they were issued provisional certificates. Later performance evaluations attested to their competence. Schools of education began graduating sufficient numbers of trained professionals to perform the jobs held by people with provisional certificates. The paraprofessionals were told they must take a series of courses in an approved program and obtain regular certification in order to secure their positions. They challenge the requirement.

A participant said that in his state, that presented no problem–that their licensure system was set up such that on-the-job evaluations, by supervisory staff, of teachers who had provisional certificates could deal with such circumstances. However, Rebell pointed out that such was not the case in several states. He noted that moreover, existence of the option did not eliminate the possibility of procedural irrationalities such as exemplified in the case of one of his clients which was outlined earlier in the discussion. Further, the thrust of the CBTE movement seems to be headed in the direction of closing off existing alternative routes to certification. He underlined this trend as a potential source of litigation under Title VII.

The scenario, in particular, has bearing upon CBTE insofar as its advo-

cates aspire to closing off alternative routes to certification other than graduation from approved programs. The likelihood of continuing teacher shortages in critical areas (a prominent current one being bilingual skills) renders such aspirations unrealistic. But more important is the observation that among the most competent practicing teachers are several who did not obtain certification through the traditional route. This fact casts doubt upon claims that the only route should be through approved CBTE training programs. This is especially true for the period in which the requirements of such programs have not been validated as related to effective teaching. However, even should there come a time when they are, it would remain an open question as to whether such programs are the *only* means to attain the competencies; and the policy question would remain as to whether or not, in the interests of fairness and quality schooling, there should be alternative routes to certification.

III. CONCLUSION

In closing the conference, it was agreed that there were, to be sure, no simple answers to the questions raised. All of the considerations voiced on various sides of the issues demonstrate the complexity of the problems involved in constructing sound, workable policies that promote quality schooling and encompass fair, rational employment practices.

An important item on the conference agenda was to elicit from the states opinions on further work we might undertake on their behalf. There was a broad consensus that a training manual for state licensing authorities be prepared that would contain an overview and summary of the status of litigation and psychometric research relevant to teacher certification and that would also contain specific model procedures for establishing certification standards for use by state administrators.

Acknowledgements

I wish to thank the U.S. Office of Education and Laurence B. DeWitt, in his capacity as Director of the Educational Policy Research Center (Syracuse Research Corporation), for supporting the production of this report. I wish to thank, also, Paul Olson for his faith in assigning me the task and Fred Baldwin, Laurence DeWitt, Emily Haynes, Nancy Jones, Michael Rebell, Maureen Webster, and David White for reviewing the draft. Unfortunately time did not permit the incorporation of their excellent suggestions for reorganization and elaboration on the context of competency-based teacher education. Moreover, any deficiencies of the report remain the responsibility of its author.

FOOTNOTES

1. Maureen MacDonald Webster, "The Quest for Better Schools," in Miriam Clasby, Maureen Webster and Naomi White, *Laws, Tests & Schooling: Changing Contexts for Educational Decision-Making*, prepared under Contract No. OEC-0-72-5017, Educational Policy Research Center, Syracuse Research Corporation (October, 1973), p. 63.

2. Phyllis Hawthorne, *Legislation by the States: Accountability and Assessment in Education*, Report No. 2 (November, 1974), Revisions, prepared for the Cooperative Accountability Project by the State Educational Accountability Repository (SEAR), Wisconsin Department of Public Instruction.

3. Max Weber, *The Protestant Ethic and the Spirit of Capitalism*, trans. Talcott Parsons (New York: Charles Scribner's Sons, 1958).

4. Sheila Huff, "Problems in Implementing Competency-Based Programs," Educational Policy Research Center, Syracuse Research Corporation, Contract No. NIE-400-75-0036 (June 30, 1975). The paper outlines some 24 elements that must be re-examined when goals of an educational institution change.

5. An excellent explication of this concept is presented in Nijole Benokraitis and Joe Feagin, "Institutional Racism: A Critical Assessment of Literature and Suggestions for Extending the Perspective," draft University of Texas, Austin (Fall, 1974).

6. EEOC Guidelines on Employment Selection Procedures, *Federal Register*, Vol. 35, No. 149, pp. 12,333-336, henceforth referred to simply as the EEOC *Guidelines*.

7. A summary of relevant portions of Title VII, together with a discussion of some of the more educationally relevant cases, may be found in Sheila Huff, "Credentialling by Tests or by Degrees: Title VII of the Civil Rights Act and *Griggs v. Duke Power Company*," *Harvard Educational Review*, Vol. 44, No. 2 (May, 1974). See also David M. White and Richard L. Francis, "Title VII and the Masters of Reality: Eliminating Credentialism from the American Labor Market," forthcoming in *Georgetown Law Journal* (Fall, 1976).

8. David Pike, "Hold the Phone on Bias, Mr. AT&T," *Washington Post* (June 10, 1976).

9. Lesley Oelsner, "Supreme Court Upholds Exam in which Blacks Scored Low," *New York Times* (June 18, 1976), p. 1.

10. *Washington v. Davis*, 44 U.S.L.W. 4789 (June 7, 1976).

11. Lesley Oelsner, *op. cit.*

12. Post-conference developments in the *Washington v. Davis* case cast some doubt upon this earlier judgment.

13. As pointed out in Michael Rebell's paper, even where it is not possible to show adverse discriminatory impact, there is a growing trend toward a revitalization of the substantive due process concept that may lead a court to seriously consider a future claim of a plaintiff who alleges that he/she has been denied substantive right to pursue a teaching career by an irrational credentialling system.

14. *Chance v. Board of Examiners*, 330 F. Supp. 203 (S.D.N.Y., 1971), aff'd 458 F. 2d 1167 (2d Cir., 1972).

15. *Rubinos v. Board of Examiners*, Civ. No. 74/2240 (S.D.N.Y. filed May 24, 1974). A decision is still pending in this case.

16. It should be noted that no one at the conference had a naive view of 'teacher effectiveness.' They realized that the notion was context bound—certain practices being appropriate in some circumstances and not in others.

17. A rare example of such studies are those constructed by Frederick J. McDonald of the Educational Testing Service, Princeton, for the Beginning Teacher Evaluation Study being conducted for the California Commission for Teacher Preparation and Licensing and sponsored by the National Institute of Education.

18. New York Board of Regents, "Statewide Plan for the Development of Post-Secondary Education," Unit 2 (1972).

OTHER BOOKS

Study Commission on Undergraduate Education and the Education of Teachers

I. Higher Education Reform

____ 1. *Toward a Community of Seekers*– transcriptions from sessions of the Johnston College National Symposium on Higher Education. An appendix includes names and addresses of people to contact at 267 experimenting college programs and related organizations. Discussants include Harold Hodgkinson, Frank Newman, Paul Dressel, and others. Basic to Higher Education reform, '75-'85.

____ 2. *The Ranking Game: The Power of the Academic Elite*– a critical look at assessments of higher education that are based on institutional reputation. Attack on the most influential rankings of higher education: Cartter, Roose-Anderson, etc. Discussion of the ways in which such reports are used and the impact of them on the homogeneity of graduate and undergraduate education. The authors are well known critics of Higher Education policy: Patrick Dolan and William Arrowsmith.

____ 3. *Nothing But Praise: Thoughts on the Ties Between Higher Education and the Federal Government*– looks at reforms in the process of granting federal money to institutions of higher education, to state agencies, and to schools. A critical appraisal of the total federal posture toward higher education and education in general.

____ 4. *Of Education and Human Community*– essays on education and the community-building process. Also includes an edited discussion among leaders in experimental education. Discussants include Harold Hodgkinson, Sam Baskins, Dwight Allen, Harvey Scribner and the like.

____ 5. *Open Admissions: The Promise and the Lie of Open Access in American Higher Education*–essays on access in the state higher education systems of Nebraska and California and the City University of New York. The book says a lot about the background of the current debate on open admission in New York City and elsewhere.

____ 6. *A Time Half-Dead at the Top: The Professional Societies and the Reform of Schooling in America 1955-1975*– an examination of 20 years of discipline-oriented "professional society" work with education reform in the schools and higher education. A strong criticism of the role of the professional societies and accrediting-licensing agencies in the shaping of higher education.

____ 7. *Forms and Formulations of Education*– several studies of how education reform is conducted in terms of various explanation-logics and the possi-

bilities and limitations of such reform viewed from the perspective of theory-construction.

____ 8. *Some New Ways of Learning and Teaching*– brief descriptions of experimental teacher education programs throughout the U.S. Includes sections on minority institutions, free schools, alternative public schools, community education, metropolitan public colleges and women's studies programs.

II. Sexism in Education

____ 9. *We'll Do It Ourselves: Combatting Sexism in Education*– an analysis of sex role channelling and stereotyping with numerous alternative action outlines. Combines both feminist and gay perspectives with blueprints for creating women and gay studies centers, legal and medical clinics and day care services. Also a resource directory.

____10. *Badges and Indicia of Slavery: Cultural Pluralism Redefined*– a look at some of the social, economic, and legal implications of pursuing a policy of cultural pluralism. Includes site visits to traditional and alternative institutions and some perspectives on sexism in relation to cultural pluralism.

____11. *Mini-Manual For a Free University*– guide for setting up autonomous, student-run education projects. Several case studies are included, with a resource guide listing hundreds of groups, individuals, projects, and publications. The resource guide includes a number of listings of women's resource material centers, centers for combatting sexism, etc.

____12. *Some New Ways of Learning and Teaching*– brief descriptions of experimental teacher education programs throughout the U.S. Includes sections on minority institutions, free schools, alternative public schools, community education, metropolitan public colleges and women's studies programs.

III. Black Education

____13. *Education By, For and About African Americans: A Profile of Several Black Community Schools*– profiles of nine independent and community-based black schools from pre-school levels to the community college level. Helpful to people who want to build their own black community-based school.

____14. *Badges and Indicia of Slavery: Cultural Pluralism Redefined*– a look at some of the social, economic, and legal implications of pursuing a policy of cultural pluralism. Includes site visits to traditional and alternative institutions. Includes site visits and reports on a number of black community-based schools.

____15. *Mini-Manual for a Free University*– guide for setting up autonomous,

student-run education projects. Several case studies are included, with a resource guide listing hundreds of groups, individuals, projects and publications. Resource directory includes listings of a number of black and pan-African organizations and resource centers.

__16. *Equality of Opportunity in Higher Education: Myth or Reality?*– by Paul Mohr. Essays by Mary Lepper, Sam Wiggins, Terry Wildman and Paul Mohr on how black colleges are surviving or being destroyed under present state plans.

__17. *Black Colleges and Equal Opportunity in Higher Education*– by Paul Mohr. An analysis of plans and lack of planning for black colleges in Mississippi, Maryland, Tennessee, Florida and elsewhere.

__18. *The University Can't Train Teachers*– a group of prominent school administrators talk about school-based training for future teachers and about how certification, accrediting, tenure and funding problems affect change in teacher education. Includes extensive useful discussions by Barbara Sizemore, former superintendent of schools in Washington and in Woodlawn, Chicago.

IV. Chicano Education

__19. *The Recruitment, Channeling, and Placement of Chicano Teachers*– a discussion and resource list related to the recruiting of Chicano teachers.

__20. *Chicano Alternative Education*– this book is divided into three sections. The first section consists of an essay "Chicanos and Schools: A Perspective for Chicano Alternative Educational Situations" by H. Homero Galicia. It presents a conceptual discussion of the shortcomings of schooling in this society and a profile of four different Chicano alternative educational efforts, with brief analytical comments on these alternatives. The second section "Alternative Chicano Educational Programs in California" by Clementina Almaguer, contains a listing and summary description of seven alternative Chicano schools. The final section contains the actual transcripts of the interviews by Clementina Almaguer with personnel from the following seven programs: Oakland Street Academy (Oakland), Case de la Raza (Berkeley), Escuela del Barrio (San Jose), Colegio de la Mission (San Francisco), Universidad de Aztlan (Fresno), Huelga School (Deland), and D-Q University (Davis). These interviews contain very useful insights into the problems and processes encountered in attempting to create alternative schools. The book contains graphics by Ralph Maradiaga, co-director of Galeria de la Raza in San Francisco.

__21. *Casa De La Raza*– subtitled "Separatism or Segregation, Chicanos in Public Education," this book is essentially a case history on the educational and legal issues surrounding the decision by the Office of Civil Rights of the U.S. Office of Education to close down Casa de la Raza. The implications of this

case are of significance to all Chicano and other ethnic-based educational programs throughout the nation. This is an extensive and thorough case history, complete with a chronology of the case, description of the program, correspondence and memoranda from the Office of Civil Rights, the Berkeley school superintendent and legal experts; a highly competent and detailed article by Susan F. Appleton is reprinted from the *California Law Review*.

__22. *Parameters of Institutional Change: Chicano Experiences in Education*—contains a section of case histories of Chicanos who attempted to produce change within educational institutions and describes limitations of those changes. Reconstructed from interviews, the eleven case histories that are presented in this book are: East Los Angeles Blowouts; A Bilingual/Bicultural Attempt in Silver City; El Paso Educational Research Project; Chicano Student Movement at San Jose; Chicano Demonstrations at New Mexico Highland University; Before Universidad de Aztlan: Ethnic Studies at Fresno State College; Memoirs of a Chicano Administrator; The Struggle for Third College at UC San Diego; Casa de la Raza; The Intrigues of Educational Vouchers: Schools and Parents; and the Alum Rock Voucher Program. Also included is a section of conceptual essays that extend the argument made by the case histories. These latter essays suggest a mode for analyzing and drawing lessons to be learned posed by the case histories presented in the first section. The essays presented are "Cultural Pluralism and the History of Urban Education: Centralization, Decentralization, and Community" CASTANEDA PANEL; "Schooling of Chicanos in a Bilingual, Culturally Relevant Context" REYNALDO FLORES MACIAS; "Cultural Neutrality in Education: Some Legal Developments and Their Possible Impact on the Future of Education" LAWRENCE FREEMAN; "To Leave to Hope or Chance: Propositions on Chicano Studies" JUAN GOMEZ-QUINONES; "Fiesta of Life: Impressions of Paulo Freire" E.A. MARES; "Education Under Capitalism—A Bibliography" CARLOS VASQUEZ.

__23. *Badges and Indicia of Slavery: Cultural Pluralism Redefined*—the ideology of cultural pluralism plus site visit reports on Chicano institutions.

__24. *Educación Alternativa*—How to organize a Chicano Alternative School by people who have done it.

__25. *Mini-Manual For A Free University*—includes listings and descriptions of a number of Chicano and Latin American resource and research centers.

V. The Education of Teachers

A. Higher education position on educating teachers:

__26. *Education for 1984 and After*—College of Education leaders discuss targeted education for teachers, cultural pluralism and other topics. Discussants include George Denemark, Vito Perrone, Alfredo Castaneda, Dean Corrigan, Joan Goldsmith, Jake Carruthers and others.

___ 27. *The Supply and Demand of Teachers and Teaching*–the present perception of teachers as "single-purpose tools" rather than as professionals who are capable of working in a variety of situations increases the disparity between supply and demand. Provides suggestions toward a model for examining manpower problems.

___ 28. *The Future: Create or Inherit*– a study of changes occurring in the schools as a result of the "knowledge explosion." What happens to teacher training when teachers begin to take on new roles in the schools and communities? Includes essays by such futurists as Harold G. Shane, Robert Bundy, Charles W. Case, W. Timothy Weaver; also conversations with Harold Garfinkel, Robert Bundy, Charles W. Case and others.

___ 29. *Manpower Newsletter*–contains manpower information data on teachers and on various states' efforts to assess and deal with needs and surpluses. Presents some teacher education manpower data models.

___ 30. *Statistics Newsletter*–contains a re-analysis of the Carnegie Commission survey data on higher education which covered 70,000 undergraduates, 60,000 faculty members and 33,000 graduate students.

___ 31. *Choices*– list of alternative and experimental programs in one state: Vermont. A model catalogue.

B. School-based positions on educating teachers:

___ 32. *What is School-Community-Based Teacher Education and Why Should School Administrators Be Interested In It?*– questions and answers about school-community-based teacher education prepared after site visits to teacher centers, schools-without-walls, and other community-oriented programs.

___ 33. *The University Can't Train Teachers*–a group of prominent school administrators talk about school-based training for future teachers and about how certification, accrediting, tenure and funding problems affect change in teacher education. Discussion leaders include Dick Foster, Bud Spillane, Barbara Sizemore, Jose Cardenas, Bob Schwartz, and others.

VI. Manpower

___ 34. *The Supply and Demand of Teachers and Teaching*–the present perception of teachers as "single-purpose tools" rather than as professionals who are capable of working in a variety of situations increases the disparity between supply and demand. Provides suggestions toward a model for examining manpower problems.

___ 35. *Adjustments of Colorado School Districts to Declining Enrollments*–statistical summaries of declining districts and several case studies

are included in this book. Introduction by Kenneth Boulding, the well-known economist. Studies by Mark Rodekohr.

__ 36. *Manpower Newsletter*– contains manpower information data on teachers and on various state's efforts to assess and deal with needs and surpluses. Presents some teacher education manpower data models.

VII. Legal Issues in Accreditation, Credentialling and Certification

__ 37. *Teacher Credentialling Reform in New York State: A Critique and A Suggestion for New Directions*– the impact of teacher credentialling laws falls far short of the ideal for which they were established. The author argues that these laws "don't protect children from incompetent and unfit teachers." He suggests that improvements in the system can be perfected, school boards should be granted the authority "to hire any college graduate with relevant subject matter preparation and to require such districts to seriously and systematically assess his performance once he is on the job."

__ 38. *Basic Legal Issues in New York State on Teacher Certification*– deals with the function of government in the professional licensing or certification of educators. The basic argument of the book is that "the right and responsibility of the state to certify teachers is a legitimate, moral, and rational use of the political power of the state, only to the extent that teacher certification protects and promotes some demonstrably legitimate public interest of the people for whose welfare and benefit state accredited schools are established."

__ 39. *Teacher Education in the United States: The Responsibility Gap*– the Study Commission's final report. Chapter topics include: (1) The Future of Schools and Children–Education of Teachers as a Community Activity, (2) Teacher Education in America Now, (3) Legal Constraints: Licensing, Accreditation, and Equity in Teacher Education, (4) Gathering Information on the Process of Educating Teachers, (5) The Federal Role in Teacher Education. Available, $3 per copy. Chapter III is the final Commission statement on licensing and accreditation in general and in teacher education areas.

VIII. Indian Education

__ 40. *Badges & Indicia of Slavery: Cultural Pluralism Redefined*–the ideology of cultural pluralism plus site visit reports on new Indian institutions.

__ 41. *Teacher Education in the United States: The Responsibility Gap*– includes consideration of special Indian concerns as regards licensing, accreditation, new sorts of institutions, and planning for the future. "This report should, and we believe will, become the basis for a profound examination of our current educational system, and its recommendations, above all, should be taken seriously." . . . WASSAJA, March, 1976.

IX. Power Structure Analysis

__ 42. *How to Research the Power Structure of Your University or College*

__ 43. *How to Research the Power Structure of Your Secondary School System*—research outlines for finding out how decisions are made in these two educational communities. These books frame questions, provide probable sources for finding answers, contain case studies, and offer bibliographies. They are intended to be used as workbooks. 1973.

__ 44. *Wheeling & Dealing in Washington: A Student's Guide to Federal Educational Politics and Resources*—this guide gives practical information on lobbying, shows studies on student financial aid and sexism in education and has a resource directory of congressional committees, federal and private agencies relating to educational issues. 1975.

These books can be ordered from:

University of Nebraska
Nebraska Curriculum Development Center
Andrews Hall
Lincoln, Nebraska 68588

All books are $2 each except *Teacher Education in the United States: The Responsibility Gap* which is $3. Please make checks payable to University of Nebraska.

How can the performance of teachers be more fairly, legally, and adequately evaluated? or *How should one evaluate teacher competence?*

How might the state best assure quality teacher preparation programs?

Sheila Huff, *Report of the Special Legal Issues Conference, December, 1975—Denver, Colorado*

The essays in this book evolved from the groundwork laid at a legal issues conference in December, 1975, sponsored by the Multi-State Consortium on Performance-Based Teacher Education. Attending that meeting were staff from state departments of education and their legal advisors, teacher union representatives, faculty of schools of education, members of the research and development community, lawyers representing the interests of parents and students—a wide spectrum of individuals with a stake in decisions made regarding accrediting and licensing issues.